

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 10:06:26 ; Search time 27321.5 Seconds

(without alignments)
684.748 Million cell updates/sec

Title: US-09-303-518D-571

Perfect score: 894

Sequence: 1 atgttcgttacaattcag.....accgctataaaccgcgtaa 894

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

GenEmbl:*

- 1: gb_da:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Length	ID	Description
1	894	100.0	894
2	807.4	90.3	897
3	807.4	154.71	897
4	807.4	90.3	897
5	799.4	89.4	897
6	799.4	89.4	897
7	791.6	88.5	894
8	347.8	38.9	369
9	98.4	11.0	215050
10	69.4	7.8	13818
11	61.8	6.9	2081
12	47.2	5.3	125020
13	45.8	5.1	201050
14	45.6	5.1	10991
15	44.8	5.0	20310
16	44.4	5.0	125020
17	43.6	4.9	1527
18	42.8	4.8	924
19	42.8	4.8	33140
20	41.8	4.7	12832
21	41.4	4.6	11291
22	41.4	4.6	9983
23	41.4	4.6	11291
24	41.4	4.6	112428
25	41.4	4.6	13881
26	41.2	4.6	204050
27	41.1	4.6	1760
28	41.1	4.6	2330
29	41.1	4.6	2996
30	41.1	4.6	40969
31	41.1	4.6	110000
32	40.2	4.5	2436
33	40.2	4.5	11532
34	40.2	4.5	13431
35	39.6	4.4	10732
36	39.6	4.4	12213
37	39.4	4.4	3570
38	39.4	4.4	202050
39	39.2	4.4	32039
40	39.2	4.4	93403
41	39.2	4.4	133814
42	39.2	4.4	10202
43	39.2	4.4	14790
44	38.8	4.3	2313
45	38.8	4.3	2331
46	38.8	4.3	3101
47	38.8	4.3	10141
48	38.6	4.3	3487
49	38.6	4.3	10027
50	38.4	4.3	2097
51	38.4	4.3	4358
52	38.4	4.3	10652
53	38.4	4.3	14053
54	38.4	4.3	40476
55	38.4	4.3	176894
56	38.4	4.3	294250
57	38.2	4.3	3974
58	38.2	4.3	11840
59	38.2	4.3	13568
60	38.2	4.3	33001
61	38.2	4.3	4494
62	38.2	4.3	84985
63	38.2	4.3	94473
64	38.2	4.3	104326
65	38.2	4.3	138039
66	38.2	4.3	340900
67	37.8	4.2	1920
68	37.8	4.2	36028
69	37.8	4.2	45370
70	37.8	4.2	347750
71	37.6	4.2	1492
72	37.6	4.2	3175
73	37.6	4.2	3196

A96538	Sequence 57
A96534	Sequence 56
AE002530	Neisseria
AX044034	Sequence 56
A96536	Sequence 56
AX081572	Neisseria
A96532	Sequence 56
AL646057	Ralstonia
AE004747	Pseudomonas
PFY14568	Homo sapi
AF429315	Ralstonia
AL646064	Ralstonia
AE001874	Neisseria
AY010120	Xanthomonas
AF429315	Homo sapi
AF118888	Bradyrhiz
AX081571	Sequence
AX067448	Sequence
AE003967	Shuttle vec
HLRGYRB	Shuttle vec
HLFINSVECA	Shuttle vec
HLFINSVECB	Shuttle vec
AP003906	Oryza sat
AP004589	Oryza sat
AL646073	Ralstonia
AB038378	Ralstonia
TAE314785	Trifolium
X61026	Wheat Glu-1
SCD6	Continuation (25 o
LMFLCHR32.24	Caulobacter
CCRGROEST	Halobacter
AE005744	Halobacter
AE005044	Halobacter
E32986	Gene encodi
AE004441	Pseudomon
RCNMF	R. capsulatu
AL646065	Ralstonia
SC7AI	Homo sapi
AC073261	Homo sapi
AC027550	Homo sapi
AE004483	Homo sapi
AE001885	Deinococc
SCB314774	Secale ce
PAHEMGEN	Secale ce
AE004442	Pseudomon
AF024589	Hordeum s
AE006938	Myobacte
BC015649	Homo sapi
AX306427	Sequence
AE004788	Pseudomon
AE004873	Pseudomon
SCC75A	Pseudomon
AL355574	Homo sapi
AP001517	Bacillus
RME132004	Rhizobium
AE008250	Agrobacte
AE009363	Agrobacte
SCC123	Streptomy
AF029343	Homo sapi
AL445288	Homo sapi
AB070940	Human DNA
AC016786	Streptomy
AB018799	Streptomy
AL442120	Streptomy
AF002998	Puau rubr
AF078693	Chlamydom
RMU19726	Rhizobium m
RHMGROESTLA	Rhizobium m

c 74	37.6	4.2	10205	1	AE007230	AE007230	Sinorhizo
c 75	37.6	4.2	119290	2	AC068449	AC068449	Homo sapi
c 76	37.6	4.2	167721	2	AC032019	AC032019	Homo sapi
c 77	37.6	4.2	168414	2	AC023146	AC023146	Homo sapi
c 78	37.6	4.2	189333	1	RME603647	RME603647	Rhizobium
c 79	37.6	4.2	208050	1	AL646083	AL646083	Ralstonia
c 80	37.6	4.2	208050	2	AC092715	AC092715	Homo sapi
c 81	37.6	4.2	270409	1	SM591785	SM591785	Alseidiella
c 82	37.6	4.2	286550	1	PSEBPOS	PSEBPOS	Pseudomonas
c 83	37.4	4.2	2910	1	AE004491	AE004491	Pseudomonas
c 84	37.4	4.2	10297	1	AE004782	AE004782	Pseudomonas
c 85	37.4	4.2	13031	1	SCGD3	SCGD3	Streptomyces
c 86	37.4	4.2	33779	1	SC9C5	SC9C5	Streptomyces
c 87	37.4	4.2	102951	2	AP003909	AP003909	Oryza sat
c 88	37.4	4.2	152588	2	AL646059	AL646059	Ralstonia
c 89	37.2	4.2	190050	1	AF240955S1	AF240955	Orthococc
c 90	37.2	4.2	373	3	AE005090	AE005090	Halobacte
c 91	37.2	4.2	13542	1	AE004989	AE004989	Halobacte
c 92	37.2	4.2	14108	1	AE005067	AE005067	Halobacte
c 93	37.2	4.2	25970	1	SC1G7	SC1G7	Streptomyces
c 94	37.2	4.2	26195	1	AL646078	AL646078	Ralstonia
c 95	37.2	4.2	203050	1	AE005067	AE005067	Halobacte
c 96	37.2	4.1	9980	1	AE002080	AE002080	Deinococc
c 97	37.2	4.1	11208	1	AE002080	AE002080	Deinococc
c 98	37.2	4.1	16911	1	SC1G7	SC1G7	Streptomyces
c 99	37.2	4.1	38900	1	MTCY63	MTCY63	Mycobacteri
c 100	36.8	4.1	1401	6	AX110939	AX110939	Sequence

ALIGNMENTS

RESULT 1		894 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	A96538				
DEFINITION	Sequence 571 from Patent WO9924578.				
ACCESSION	A96538				
VERSION	A96538.1				
KEYWORDS	GI:6780161				
SOURCE	unclassified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 894)				
AUTHORS	Pizza, M., Scariato, V., Rappuoli, R., Grandi, G. and Masignani, V.				
TITLE	Neisserial antigens				
JOURNAL	Patent: WO 9924578-A 571 20-MAY-1999;				
	PIZZA MARIAGRAZIA (IT); SCARILATO VINCENZO (IT); RAPPUOLI RINO (IT);				
	CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)				
FEATURES	Location/Qualifiers				
source	1..894				
	/organism="unclassified"				
	/db_xref="taxon:32644"				
BASE COUNT	216 a 270 c 233 g 175 t				
ORIGIN					

Query Match 100.0%; Score 894; DB 6; Length 894;
Best Local Similarity 100.0%; Pred. No. 3.7e-198;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	atgttcgttaacatcagcgtgttccccccttcgcaacgagcgaatcgcgtgtg	60
DB	1	ATGTTTGTTTACAAATTCAGGCTGTTCCTCCCTTCGGAACCGCCATGCACATCCTGTTG	60
QY	61	acgcgcctgtcctaagcctccctccctcgtcgtcgttcgttcgttcgacacgctgggaac	120
DB	61	ACCGCCCTGCTCAATATGCTCTTCCTCTGCTGCTTCCTGCAACAGCGTGGGAAC	120
QY	121	cgagctcgacatctcgcgttttacccttcaaggaagacgagcgagcgtcgcgaat	180
DB	121	CGGCTCGGACATCTGCGCTTTTACCTTTAAAGAGACCGCGCGGCAATCGTCCCAAT	180
QY	181	atgcgcagcgagcggtttgaaccccgacacgacgacgacgacgacgacgacgacgacg	240

DB	181	ATGCGCGAGCGGGGTTTGAAACCCGACACGAGCGGTCAAAACCCGTTTGGGGAACG	240
QY	241	gaaatcgcgttttgaacatcgtcccgcggttttcaaaaacggaagacatcgaaaca	300
DB	241	GGAAATGCGGTTTGAACTTGCCCGCGCTTTTCAAAAACCGGAACATCGGAACA	300
QY	301	atgttcgaagcgttgcagcgttcggaacacgtgcagacgcttgcgaagcggaag	360
DB	301	ATGTTTCAAGCGGTACAGCGCTGGGAACACGTGCAGCAGGCTTTGACAAAGCGGAAGG	360
QY	361	ctgcgtttatcaacgcgcacatcgcgcagctagatttggcggaacgttcacatcag	420
DB	361	CTGCTTTTATCAGCGCGGACATCGCAGCTACGATTTTGGCGGAGCGTATACGACAG	420
QY	421	cagcttcgcttccacgtcgcgcacatcgtacacgcccgcgaatacaagcgatagaca	480
DB	421	CAGCTTCCGCTTCACCTGACCGCATGTCATCAAGCCGCGGAATCAAAAGCATGACAAA	480
QY	481	atcatgcaagcgagcgaggtgcgcgcaaaagcaaacgcgcgcgcgcgcgcgcgcgcgc	540
DB	481	ATCATGCAAGCGGAGGAGGTCGCAAAAGCAAAACCGCCGATGCAATGCAAGG	540
QY	541	gcaaacaaatcaatcagcgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	600
DB	541	GTCMAACAAATCATCAAGGCGCTGCGCGCGCGGCGAGACCATCATCTGCGCCGAC	600
QY	601	gtcccttcctccgcaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag	660
DB	601	GTCCTTCTCTCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	660
QY	661	accatgacactgcgcgcgaatattgcacacgttcaaaagcgctgaacccctgtttctgc	720
DB	661	ACCATGACACTGCGGCGGCAAAATGCGCACACGTCAAAAGCGGTGAACCCCTGTTTCTGC	720
QY	721	tgcaacgcctgc	780
DB	721	TGCAACGCGCTGCCGCGAGCAAGCGCTGCTGTCGACATTCGCCCGCTCCAGGAGGA	780
QY	781	ttgaacggaacaaagccacatgc	840
DB	781	TTGAACGGCAACAAAGCCACGATGCGCGCTGTCACACGCAATCCGCAATTTGGATA	840
QY	841	cgccgtttccgacgacatctgtttatgtacaaacgcgtataaagcgcgata	894
DB	841	CGCGTTTCCGACGACATCTGTTATGTACAAACGCTATAAAGCGCGTAA	894

RESULT 2		897 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	A96534				
DEFINITION	Sequence 567 from Patent WO9924578.				
ACCESSION	A96534				
VERSION	A96534.1				
KEYWORDS	GI:6780159				
SOURCE	unclassified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 897)				
AUTHORS	Pizza, M., Scariato, V., Rappuoli, R., Grandi, G. and Masignani, V.				
TITLE	Neisserial antigens				
JOURNAL	Patent: WO 9924578-A 567 20-MAY-1999;				
	PIZZA MARIAGRAZIA (IT); SCARILATO VINCENZO (IT); RAPPUOLI RINO (IT);				
	CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)				
FEATURES	Location/Qualifiers				
source	1..897				
	/organism="unclassified"				
	/db_xref="taxon:32644"				
BASE COUNT	223 a 266 c 227 g 181 t				
ORIGIN					
Query Match	90.3%; Score 807.4; DB 6; Length 897;				

Accession	Conservative	Mismatches	Indels	Gaps
U040	0	46	3	3

1

[illegible]

[illegible]

Mon Jul 1 09:26:19 2002

us-09-303-518d-571.rge

RESULT 11
PFY14568/C

DEFINITION 2081 bp DNA linear BCT 08-JUL-1998
Pseudomonas fluorescens tag gene and partial glyo, htrb genes.
ACCESSION Y14568.1 GI:2330640
VERSION glyo1-trna synthetase; glyo gene; htrb gene; tag gene.
KEYWORDS Pseudomonas fluorescens.
SOURCE Pseudomonas fluorescens
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.

REFERENCE 1 (bases 1 to 2081)
AUTHORS Dekkers, L.C., van der Bij, A.J., Mulders, I.H., Pheleich, C.C.,
Wentwood, R.A., Glandorf, D.C., Wilfelm, C.A. and Lugtenberg, B.J.
TITLE Role of the O-antigen of lipopolysaccharide, and possible roles of
growth rate and of NADH:ubiquinone oxidoreductase (nuo) in
MCS365 competitive tomato root-tip colonization by Pseudomonas fluorescens
98340543

JOURNAL Plant Microbe Interact. 11 (8), 763-771 (1998)
MEDLINE 2 (bases 1 to 2081)
REFERENCE Dekkers, L.C.
AUTHORS Direct Submission
JOURNAL Submitted (11-AUG-1997) L.C. Dekkers, Institute of Molecular Plant
Sciences, Department of Plant-Microbe Interactions, Wageningen
64, 2333 AL Leiden, NETHERLANDS
Location/Qualifiers
1. 2081

FEATURES
Source

gene /organism="Pseudomonas fluorescens"
/strain="WCS365"
/db_xref="taxon:294"
complement(1..758)
/gene="htrb"
complement(<1..758)
/gene="htrb"
/codon_start=1
/transl_table=1
/protein_id="CA74897.1"
/db_xref="GI:2330641"
/translation="MDKLGALLVGLALFLALLRAVAVAGSAGMFMKLPNSRD
VRLNLAKEPPEPAERERLVSGLKIGSLKSLSCAMIMPAOSIDIVEVELE
GKAVASTKEGILVSHLGNVETVNEKSGQKPIFFYPPRLKAVDELKORVOL
AGKAVGFLHMLRLPDGSKYKVLLEAPAMYST"

gene /gene="tag"
complement(802..1359)
/gene="tag"
/codon_start=1
/transl_table=1
/protein_id="CA74898.1"
/db_xref="GI:2330642"
/translation="MPRCMSDDPLVYAVYDQMGTPRLDAGLPELILLEGFOAGL
SWIVLAKRERYREVLVYGFQDQVORVAKMSDAIDLDMDPGIIRNRLKAAAKRQAW
LALEDVPLMSEFVDKPLINFKRTEVPAITPEALAMSKLAKGFTFVPTICYA
1441..2081
/gene="glyo"
1441..2081
/gene="glyo"
/codon_start=1
/transl_table=1
/product="glyo1 trna-synthetase"
/protein_id="CA74899.1"
/db_xref="GI:2330643"
/translation="MSQPTPAVTRPDIILALQGYMAEGCQVYLQPYDMVAGTFFHT
ATLRAIGETWNAAYVOPSDRGTGENPNRLQHYQFOVVLKPNENFQELVIG
SLKRVGLDELVDHIFVEDNNEPFLGAWLGMGWEVLMNGEIVQFTFOAGGICICYP

gene
CDS

BASE COUNT 406 a 643 c 632 g 400 t
ORIGIN VGGTITGIERLAMYLOCVDSVYDLWADDPFGKTYGVDFHONVEOSTL

Query Match
Best Local Similarity 47.2%
Matches 263; Conservative 0; Mismatches 282; Indels 12; Gaps 2;

QY 234 ggaacgcgaataatgcggttgtaactgcccgcgttttcaaaaacgcgaagcat 293
DB 546 GCACATGCGCAAAATCCCTGACCAAGACCCCTGCTGATCTGCGCCGACGCTC 487
QY 294 cgaacaatgtcaagcggtgtaacggtgtaacggtgtaacggtgtaacggtg 353
DB 486 CATGACCTGCTGCTGCAAGTCAAGGCTGAGGTGCTGAAAGACGCTGCTCCG 427
QY 354 cgaagcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 413
DB 426 CAAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
QY 414 cagcagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 473
DB 366 TTGCAAGCA-----GTGCAACCGATATTCTTCCGCGCCCAAGCGGT 313
QY 474 agcaacaatgtcagcgcggtgctgctgctgctgctgctgctgctgctgctgctg 533
DB 312 GGATGAATTGCTGCGCAACACGCGCTGCACTGCGGCAAAATGCTGCACTGCT 253
QY 534 acaaggggtcaacaataatcaagaagcctgctgctgctgctgctgctgctgctgct 593
DB 252 GGAAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
QY 594 cgaacgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 653
DB 192 CCA-----CCGGAACCGCGCAATCCCGGTATCTTCTGCTTCTTCCACCA 139
QY 654 tgcataccatgacacgtgctgctgctgctgctgctgctgctgctgctgctgctg 713
DB 138 GCGCTTCCGCAAGTTCGTACCGAATGCTCCCGCGGCAAGGCGTCTTCCACCA 79
QY 714 ttctgctgtaagcgtgctgctgctgctgctgctgctgctgctgctgctgctg 773
DB 78 CCGTCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79
QY 774 agcggaattgaacgca 790
DB 18 GAGCGCGATGTACACA 2

RESULT 12

LOCUS AF429315 125020 bp DNA linear PRI 18-JAN-2002
DEFINITION Homo sapiens junctional 3 (JPH3) gene, partial cds.
ACCESSION AF429315 GI:17646244
VERSION
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 125020)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Holmes, S.E., O'Heare, E., Rosenblatt, A., Callahan, C., Hwang, H.S.,
Ingber, N.T., Ross, C.A., Fleisher, A., Stevanin, G., Brice, A.,
A repeat expansion in the gene encoding junctional 3 is
associated with Huntington disease-like 2

JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE 2 (bases 1 to 125020)
REFERENCE Holmes, S.E., Ingber, N.T., Ross, C.A. and Margolis, R.L.
AUTHORS Direct Submission

TITLE

Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
Location/Qualifiers

```

/chromosome="16"
/chr_pos="16q24.3; between D16S520 and WI-12410"
/map="16q24.3; between D16S520 and WI-12410"
/notes="isolated from a patient with Huntington's  
Disease-Like 2 (HDL2)"
Disease-Like 2 (HDL2)
complement(35581..35746)
/rpc_type=tandem
repeat_region

```

```
mRNA      /rplL-uncat(<36507..>.36887)
           complement(<36507..>.36887)
gene       /gpe="jph3"
           /product="junctional phl in 3"
           complement(<36507..>.36887)
CDS        /gene="jph3"
           /note="jph3"
           complement(<36507..>.36887)
           /gene="jPH3"
           complement(<36507..>.36887)
```

ORIGIN

Query Match	5.3%	Score 47.2	DB 9	Length 125020	
Best Local Similarity	10.2%	Pred. No. 1.1	261	Indels 7	Gaps 3
Matches 68	Conservative 330	Mismatches 103			

TITLE	Genom
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 201050)
AUTHORS	Boucher C. A.
TITLE	Direct Submission
JOURNAL	Submitted (05-DEC-2001) genome and CNRS UMR-8030, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS, BP27, 91326 Castanet-Tolosan Cedex, France, Fondation Jean Dausset-CEPR, 27 rue Juliette Dodu, 75010 Paris, France, LMCM CNRS 118 Route de Gaston Cremieux, CP5706, 91057 Evry Cedex, France, URGV, 2 rue Gaston Cremieux et Intelligence Artificielle INRA, BP27, Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27, F31326 Castanet-Tolosan Cedex, France, Laboratoire de INRA, BP27, F31326 Castanet-Tolosan Cedex Christian.Boucher@toulouse.inra.fr/R.solanacearum.html. http://sequence.toulouse.inra.fr/R.solanacearum.html. Location/Qualifiers
COMMENT	
FEATURES	1..201050
SOURCE	/organism="Ralstonia solanacearum"

```

FEATURES
    source
        1..201050
            /organism="Ralstonia solanacearum"
            /strain="GM11000"
            /db_xref="taxon:305"
    206..1447
        /gene="RSC1326"
        /note="RS02848"
        /nc="1447"
    206..1447
        /gene="RSC1326"
        /EC_number="7.6.1.1"
        /function="small molecule metabolism; amino acid
            biosynthesis"
        /note="Product confidence : probable
            Gene name confidence : hypothetical
            predicted by Homology
            predicted by Framed"
        /codon_start=1
        /transl_table=11
        /evidence=not_experimental
        /product="PROBABE ASPARTATE AMINOTRANSFERASE PROTEIN"
        /protein_id="CAD15028.1"

```

```
QY      581 ccalcatcctgcggccagcacgltcccttccttcgcgaagaagcgcg---gcgtgtggcg   636  
       ::::::::::::::|:::|:::|:::|:::|:::|:::|  
Db    51112 KGCSTATKSMAMTYSCMRKRKTGRSSAWMYMCCSKARASGMYCNMMKGCGYYGGGKG 51171  
QY      637 gatcttctcgcaaacctgcataaccatgcacattgcyggccaatltagcacacgtlcana 696  
       :||:::|:::|:::|:::|:::|:::|  
Db    51172 SASITWMYMCCSRACKYTTTRNSSGDRGNKATPWTFTTRRSSRMSWITYKANMWRBNTWG 51231  
QY      697 ggcagtg 702  
       ::::|:  
Db    51232 KKMKTK 51237
```

RESULT	13		DNA	linear	BCT-01
AL646064		201050 bp		complete sequence;	
LOCUS	Ralstonia solanacearum GMI1000 chromosome,				
DEFINITION	Ralstonia solanacearum segment 8/19.				
ACCESSION	AL646064 AL646052				
VERSION	AL646064.1 GI:17428340				
KEYWORDS					
SOURCE	Ralstonia solanacearum.				
ORGANISM	Ralstonia solanacearum; beta subdivision; Ralstonia group; Bacteria; proteobacteria;				

REFERENCE AUTHORS

TITLE Genomes of *Arabidopsis thaliana* accessions Col-0 and Ler-0
JOURNAL unpublished
REFERENCE 2 (bases 1 to 201050)
AUTHORS Boucher, C. A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
des Plantes-Microorganismes INRA-CNRS, 171 Avenue de la Station Jean
Baptiste Lamarque, 91190 Brunoy, France

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
Boucher, C.A. Direct Submission Submitted (05-DEC-2001) CP5706, 91057 Evry Cedex, Cremieux, Castanet-Tolosan BP27, 31326 Castanet-Tolosan Dausset-CERH, Narbonne, URGV, 2 rue Gaston Cremieux F31326 Castanet-Tolosan Cedex	genoscope and CNRS UMR-8030, 2 rue Gaston Laboratoire de Biologie INRA-CNRS, Micoorganismes Jean Paris, France, LMCM CNRS 75010 Paris, France, genoscope and INRA Evry Cedex, France, genoscope et Intelligence Artificielle INRA, BP27, Laboratoire de Genetique Cellulaire Castanet-Tolosan Cedex
Christian.Bouchet@toulouse.inra.fr http://sequence.toulouse.inra.fr/R.solanacearum.html .	

FEATURES
source

gene

CDS

/db_xref="GI:17428341"
/translation="MKTIQKSAKLNNVCYDIRGPVLEKAKOMEKEGHQIILKINGLA
PGFDAEEIQDMIRLNPNSAGYSKSGIFFAKRAVMHYTOOGIKNTLIDYIGN
GASLLATLNLADGDELPPADYPLMTMTSLSGGPHVYLCDESNKWPDLDD
IRAKITPNTKGIYVNPNNPTGALYSDALLHDLIALAREHGLVITFADEVYDVLFDN
RHTMAASLSDVLTVPFNSLSKSYRSCGTRAGMVAVSGKRPAPKDIIEGLNMSRL
CAVPGQMAIQTLAGGQSIINDIAPGGRMRORDLAYELITAIIPGVSCVKPKAALYM
FPRIDPSVPIDDQTFIRQLLEERVLVQGTGFNMHSPDHFRIVELPHEDDLREAI
GRIARFLERYRORHGTGIRAA"
1538..2857
/gene="RSC1327"
/note="RS02849"
1538..2857
/gene="RSC1327"
/EC_number="1.1.1.3"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE HOMOSERINE DEHYDROGENASE OXIDOREDUCTASE
PROTEIN"
/protein_id="CAD15029.1"
/db_xref="GI:17428342"
/translation="MKPIKIGLIGLCTVGGGTLKYLQNRQEIERRAGIGIEVALIAV
KNVRAKAIQVETGASVPVDDPVAVADSPEDIDVIELIGDYDLRELVRAIANGKH
VYANKALLAHQNEIFKAADKGVWVAPEADIGIPILKILREGLTNRLEWJAGI
INQTFNFISEMDEKGLDFGVLAEOKRGVAADPTIEGFDAAHKLTLLASIAFG
VPQFRAVVEGTTKLEAVDIRAEELGRIKILGTRAEAGIELRVPPTVPATRL
IANVEGAMNAAVVGDAVGPITLYCGAGAPETASAVVADIVYTRLRHDEPHVPH
LAOPDALSNPLPIIEYRSAYILNMRVADETVGLADITRLAESGISIDMLQKES
RSEGPOTDITIMSHVVEKOVNNAIATIALPTVLSKTRLMEEILN"
2889..4334
/gene="thrc"
/note="RSC1328; RS02850"
2889..4334
/gene="thrc"
/EC_number="4.2.99.2"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE THREONINE SYNTHASE PROTEIN"
/protein_id="CAD15030.1"
/db_xref="GI:17428343"
/translation="MQYRSTRGHAEPQSFSLILGLAPDGGLYLPEQYPOVSAEELT
RMRDLPLADLAELIRKATDIPEDDLALTRRTYTPVEYCNARNGDMADITPLHRL
GEAGTSLILCSNGPTLAFKQMAOLGDLFEVALDREHAEILNLTGATSGDSSAA
EYAMGKRGVREMLSPHRKMSAFOTAOEFLSLODKININIAIEGVFDNAODIVKASN
DLAEYAKOKITGVNSINARVAVVYVYFKYLLAPKIGDKVSCVSGNFGNACAG
HYFDMGLPIDKLVYATNENVDLDFEFTRTGTRVSAQVTHITSPSMDISKASNEER
EYFDLGRDPAKLAQLFNDVDEKGSMTGKREFDRIAEFGVSGRSTHEDVNTIRD
VVRGTGVNVDHTAGVAVRENILVGVPMVVLFTVALPAKFGDTIRALDREPERPAG
VADISLPQREFVMPADADRIRKXYIAQHTGL"
4361..4957
/gene="RSC1329"
/note="RS02851"
4361..4957
/gene="RSC1329"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical

predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD15031.1"
/db_xref="GI:17428344"
/translation="MAEAAFSCLHDIRACRLCGALPEHPRPVVLADRRARVILVQ
APGRIVHETGIPMNDRSGDRLEWMLRIGYDAFYADRIAVVPMGFCFPGTSPNDLPP
RPACAPQWRMLPLTAMPDLKTLALVGYAQRVYLDVPRGMPLTEFVVRGVPRADEGAV
LEPLPHSPBNRLAMFKHNPWFDAQIVPRLRALAAAFV"
5037..6296
/gene="moa2"
/note="RSC1330; RS02852"
5037..6296
/gene="moa2"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/note="Product confidence : putative
Gene name confidence : putative
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS MOEA
PROTEIN"
/protein_id="CAD15032.1"
/db_xref="GI:17428345"
/translation="MSANDAPAVPSSLITLALLOLLAARCLTETETVPTLHANGR
VLAQVNRGLVPPADNVTQMDGVAVRADITAPGTRKVAQRIPAGVGRPLAAGEAA
RIFGALLIPGADAVVMQEQKPPDGTIVLIDVPEAGEMIRAGDITAGAEILPR
GTRGAQDLGLAASVGCAGLVRRPVAVEFTDELAAMGEPLPGAIIKNSRFTLR
ALLNLTGCEITDFGIVPTLQATBDTIREAAEADLLITSGSVSGEEDHIRKAVAE
GRUNIMQAIKRGKRLAEGVRRPGMAGGPAFFIGLPGNPVSSFTFTLLFVRPFL
RLGVDVPTPKRIMRADFALPKDGRGNAEFLRANINNGGGLDLPNOSSGVILSTWVG
DGLIDNPNPQPIAAGDVTAFPLPAGLV"
6310..6573
/gene="moaD"
/note="RSC1331; RS02853"
6310..6573
/gene="moaD"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE MOLYBDOPTERIN MPT CONVERTING FACTOR
(SUBUNIT 1) PROTEIN"
/protein_id="CAD15033.1"
/db_xref="GI:17428346"
/translation="MDMQIDIRFFASVREQLGTSHEAAVAVDTVTVGLRLRWLAAG
PAWETLAEGRALMALHDHVVYADAPLRGCEVAFFPVTGG"
6586..7116
/gene="moaE"
/note="RSC1332; RS02854"
6586..7116
/gene="moaE"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/note="Product confidence : probable
Gene name confidence : probable

Query Match 5.1%; Score 45.8; DB 1; Length 201050;
Best Local Similarity 50.2%; Pred. No. 2.3;

	Matches 113; Conservative	0; Mismatches 112; Indels	0; Gaps	0;
OY 348	caagggcgaagggctgctgttcatcaagccgcacatgcygcagctciagattggcgagc	407		
Db 12263	CATCGACGAGCTGTGTGTTCACCGCCTCGACCAAGGGCCCAATCGATGATCGGGG	12322		
OY 408	ctaatcaagcagcagcgtctccgtctccactgacgcgcataagcgcgcgcgaatactaa	467		
Db 12323	CATTCAGCTGCACAGGGCTTGGCGGGCTTGGCGGACATGAGACTGAACTGAACTGGAGATCAAG	12382		
OY 468	agcgatagacaaatlcatgcaagcgagggtgcgcggaagcgaataacgcgcgcac	527		
Db 12383	CGACCCGGCGTCGCCAACCTGGCTCGCGGGGGTACACCCGGTGTTCCGGCGCGGGCC	12442		
OY 528	cgggatcacaaggggttcaaaccaatcatcaagagccctgtgcgcgg	572		
Db 12443	GCTTCAGCGCGCATTCACGACGATTCGAGAACCCGGTCCGGG	12487		

RESULT	14
AE001874/c	
LOCUS	10991 bp
AE001874	DNA, linear
DEFINITION	Deinococcus radiodurans R1 section 11 of 229 of the complete

ACCESSION	AE001874	AE000513
VERSION	AE001874.1	GI:6457778
KEYWORDS		
SOURCE		
ORGANISM	Deinococcus radiodurans.	Deinococcus radiodurans.

ORGANISM	REFERENCE
<i>Deinococcus radiodurans</i>	White, O., Eisen, J. A., Hickey, E. K., Peterson, J. D.,
Bacteria: <i>Thermus</i> / <i>Deinococcus</i> group; <i>Deinococcales</i> ; <i>Deinococcaceae</i> ;	
<i>Deinococcus</i> .	
1 (bases 1 to 10931)	

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (Passes 1 to 10991)	White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Podson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffitt, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M., et al.	Genome sequence of the radioreistant bacterium <i>Deinococcus radiodurans</i> R1	Science 286 (3444), 1571-1577 (1999)

REFERENCE	2 (bases 1 to 10991)
AUTHORS	White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., McFie, K. S., Qin, H., Jiang, L., Pamphile, R., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, D., Utterback, T., Zalewski, C., Makarova, K. S., Aravind, L., Daly, M. J., Minton, K. W., Fleischmann, R. D., Ketchum, K. A., Nelson, K. E., Salzberg, S., Smith, H. O., Venter, J. C. and Fraser, C. M.
TITLE	Submitted Submission
JOURNAL	Submitted (08-NOV-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
SOURCE	1..10991

```

/strain="R1"
/bb_xref="taxon:1299"
/chromosome="1"
/complement(131. .1015)
/gene="DR0112"
/complement(131. .1015)
/gene="DR0112"
/notes="similar to GP:3342798 percent identity: 55.95;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamine cyclotransferase"
/protein_id="AAFP09704.1"
/db_xref="GI:6457779"
/translation="MKRSPALSLTFEALIALAIPLASAPPTBPAPAKVSARTI
VLPVAAPIHPRRAAFATGGLDIYGGGHIESTQGVESPLRNSELKGAKYIWMSPIL
QALPAQFEGSGTOLGSIYQLTWQDGAULTIDATFETETRNHIOEGMGKILSDGSG
INSNGIYVWMDPPTFAAQRSSVOYTDGQGVFNRLNLELYVGGSYIANWGLTIDLRAS

```

```

gene
CDS
    HPOTGKGVGLWIDVSLTTEVSAATKKGQALTFEDDPVNGIAIFPERGTLTLTGKRPET
    LPEVKKPGGLGKLVKVEGAKN"
    complement(1098. .2120)
    /gene="DR0113"
    complement(1098. .2120)
    /gene="DR0113"
    /note="similar to SP:P32573 P1D:1302205 P1D:553143
    P1D:600063 percent identity: 62.84; identified by sequence
    similarity; putative"
    /codon_start=1
    /transl_table=11
    /product="oxidoreductase, short-chain
    dehydrogenase/reductase family"
    /protein_id="AAFO9705.1"
    /db_xref="GI:6457780"
    /translation="MDWGKGRPKFLCKRFLFCIGRSYVLRRLFGIGRCSYLSGRFSI
    RPHVYPTPRFESMTQSSPNRPADPATTFTRFDLLAGKRALITGGSGINIGLQSPAAH
    GCAVTLINPELLEKQAQNAAGQIVDAGGRAMGSAVDYRFALEAAVAATAEIGGFIDV
    LAAQANIPGAPVDGISPNIGFTYVIDLTGTYNTIKAAAPRLKVGNGTSLISAYGVA
    VPMQAIHVAAKGVDALITQTLAVEMGLRGVRVANA.IIGP.IDGTEGMARLADPEKSRSA
    FTTTVLGRFGVPQDIIANALFLVSDASIVTYGVLIPVGGQNMGLGAPQGYOMTLQMG
    AAQOKD"
    complement(2344. .3117)
    /gene="DR0114"
    complement(2344. .3117)
    /genes="np0114"

```

```

/note="similar to PID:J502224 GB:A123456 percent
identity: 55.46; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="penoyl-CoA hydratase, putative"
/protein_id="AA09706.1"
/db_xref="GI:6457781"
/translation="MYOSIRLTQRPUSDGTTGGTVAITLFLAAMSGSPAFWQOE
POLSEIGDARVILYRGQVFSAGIDVSNCAALVPLGPDFAKAVDEMAVTEGI
AALPMPIYIAVHGWCIGAGTELIGADRLCSDARPSPLEYLTGTLGLOLPL
LIGSGRIHALGEAIDATIRKGLVTEVLPDQALPRRABALAHNLALPKALE
GTRALSDGLPAHSEIAAAVRRNAEHMTVEALGMRKG"
gene
3230..4120
/gene="Dr0115"
3230..4120
CDS
1

```

gene
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA09711.1"
/db_xref="GI:6457786"
/translation="MGLVYCRRLRALRFNMAPEHGMGLRSGAAPGSGMKYRALPLP
VSVASLLAALGGAQPPICITTPVPVOTAPAPAPVACTOGAMAKKAIDIVYDKGIG
FIPGDSPEWCSAITROEVAQVLAIRLLAMPENTFNPALDITLRGVAELAGLEEIE
RAGLAQNGSIDIELRAQIDELRTALNNPAGGCEAGALRPVRAKPAAPSVAPRAK
LRVLAAPVIRKIVPSVPAPRAVSAARVITPFRPSPSATATWAPPTVASSRTV
ARMSAPPATRTSTASACA"
4018..4407
/gene="DR0116"
4018..4407
/gene="DR0116"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA09712.1"
/db_xref="GI:6457787"
/translation="MGAAYSIILQNGWQWVRITAGNDQIYGFGRVATGDVALGETE
PGNSLSAVYRGTGTRPFGCILGAGAGYFNMASTPELSIGDYIVVDRAVIAFMAEA
QHYFENINPNRNKSLSTAAGFKVR"
complement(4521..5858)
/gene="DR0117"
complement(4521..5858)
/gene="DR0117"
/note="similar to GB:LL4862 SP:Q06862 PID:289133 percent

Query Match	Best Local Similarity	Score	DB 1;	Length	10991;
Matches	165;	Conservative	0;	Mismatches	199;
				Indels	0;
				Gaps	0
QY 274	ltcaaaaacccgaagacatcgaacaagtccaagcgtacacgctggaacacgtg	333			
Db 7146	TTCCGGAACATGGAACCGCTCGATTGTGGACCGACCTTCAACAGACCGAAACTTG	7087			
QY 334	cagcaagccttgacaagggcgaagggctgctgttcatacgcgcacatcggaactac	393			
Db 7086	CAGGACATCTACGTCGAAGGCAAGAAAGTACGATCCTGTACCCCGACGAGTAC	7027			
QY 394	gatttggcggaagctcatalcaagccgaagcttcgttcacactgaagcgaatagaag	453			
Db 7026	GTGTTATTGAGCATGGAACCTTCGACGACGGTTCACTTCGGCAGAACATGTGATGCGAC	6367			
QY 454	ccgcgcgaataatcaagacgatatagaacaatcatgcagcgcggaaggtgctgcgcaagcg	513			
Db 6966	GCCGCCAAGTTTCTATGAAGGAAACACCGAGTCTGGAAGTGGGATGTACGGCGCAGACGCC	6907			
QY 514	aaacccgcgcacccacgcgcatcaaggggtgcaacaacatcatcaagggccctgcgcgagcg	573			
Db 6906	CTGAGCATCAAGCCTGCCCAACGATCTCTCAAGATCACCAGACCGACCCCGCGGTG	6847			
QY 574	gaagcaacatcatcctcgtccgcgaacgctcccttcgcgcaagggcgagcgctgtg	633			
Db 6846	CGCGCGCACACCGTGTGGCGCGCACCAAGCCCGCACCTGTAAGAACCGGCGCGTGTG	6787			
QY 634	gcgcg	637			
Db 6786	CAGG 6783				
RESULT 15	AY010120	20310 bp	DNA	linear	BCT 29-NOV-2001
LOCUS	Xanthomonas oryzae pv. oryzae xanthomonadin biosynthetic gene				
DEFINITION	cluster, partial sequence.				
ACCESSION	AY010120				
VERSION	AY010120.2	GI:15055553			
KEYWORDS	Xanthomonas oryzae pv. oryzae.				
SOURCE	Xanthomonas oryzae pv. oryzae				
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;				
REFERENCE	1 (bases 1 to 20310)				
AUTHORS	Goel,A.K., Rajagopal,L., Nagesh.N. and Sonti,R.V.				
TITLE	Xanthomonadin biosynthetic gene cluster from Xanthomonas oryzae pv. oryzae				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 15118)				
AUTHORS	Goel,A.K., Rajagopal,L., Nagesh.N. and Sonti,R.V.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-OCT-2000) C.C.M.B., Uppal Road, Hyderabad, A.P. 500 007, India				
REFERENCE	3 (bases 1 to 20310)				
AUTHORS	Goel,A.K., Rajagopal,L., Nagesh.N. and Sonti,R.V.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-AUG-2001) C.C.M.B., Habsiguda, Hyderabad, A.P. 500 007, India				
REMARK	Sequence update by submitter				
COMMENT	On Aug 1, 2001 this sequence version replaced gi:11693112.				
FEATURES	location/qualifiers				
source	1..20310				
	/organism="Xanthomonas oryzae pv. oryzae"				
	/db_xref="taxon:64187"				
	/note="pathovar: oryzae"				
	218..1999				
CDS	/note="orf1: similar to YTFM from Escherichia coli"				
	/codon_start=1				

OY	555	caaggcccttgccgagggcgaagcaacatcatctgccgacacgcttccttcgca	614
		: : : : :	
Db	17261	MAGSCTTGGSSWNNRRNRNRMGKTGCNTMYRRSRAMNNGNAALAGCTTCCCATNTNGG	17202
OY	615	ggaagcgcgcgcgctgtgtggcggatlttttcggcaacctgcatacaccatgcac	674
		: : : : :	
Db	17201	GGAAAAAGGGCCSASARASCYKGRMSSKCYRGTRRRCKMSKCRGSRNKSMMGRG	17142
OY	675	ggcaaatgtgcacacgctcaaaagcgctgaaacctgtttctgtcgtgcagcgcctgc	734
		: : : : :	
Db	17141	GKTSYSAKSGSGGCTCYCWMWGRKGRKCMSRRMKMYKSYRRKRMWTCKMKCYSYMMMY	17082
OY	735	cgacgagcaaggttcgtgtgttcacatccgcgccgctccaagyggaattgaa	794
		: : : : :	
Db	17081	CRSMCCMCMCKKCCGCCYSMGMSYSYSGKYSMGMSYMMRSYSKRSTSKARSSKRCGG	17022
OY	795	aagccacgatacgccgcgctgttcaaccgcacatccgaatatgttaacgcgcttlttcgac	854
		: : : : :	
Db	17021	TGGRYKGGGRSYGKGGGSYKGGGSMWGKGKGRSTRSSAGKSKSYKMSCCARYKSKSYSK	16962
OY	855	gcagatcatctgttatgtataaac	877
		: : : : :	
Db	16961	CYSRRMCMCSYCYCWGGKCMSSC	16939

FEATURES	source
LOCUS	AF118888
DEFINITION	Bradyrhizobium japonicum malonyl CoA synthetase gene, partial cds.
ACCESSION	AF118888
VERSION	AF118888.1
KEYWORDS	GI:6640944
SOURCE	Bradyrhizobium japonicum.
ORGANISM	Bradyrhizobium japonicum.
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group.
AUTHORS	Bradyrhizobium group.
TITLE	1 (bases 1 to 1527)
	Koo, H.M. and Kim, Y.S.
	Identification of active-site residues in Bradyrhizobium japonicum
	malonyl-coenzyme A synthetase
JOURNAL	Arch. Biochem. Biophys. 378 (1), 167-174 (2000).
MEDLINE	20326872
PUBMED	10871057
REFERENCE	2 (bases 1 to 1527)
AUTHORS	Koo, H.M. and Kim, Y.S.
TITLE	Direct Submission
JOURNAL	Submitted (08-JAN-1999) Biochemistry, Yonsei University, 134
	Shinchon-dong, Seodaemun-gu, Seoul 120-749, Korea
	Location/Qualifiers
	1..1527

```

/strain="USDALI10"
/db_xref="taxon:375"
1. >1527
/codon_start=1
/transl_table=11
/product="malonyl CoA synthetase"
/protein_id="AAF28840.1"
/db_xref="GI:6840945"
/translaction="MNRKANANLFSRLFDGLDDPKRIALIEHDGARISGDIILAAAGC
MANVILVARGCPDGRVAQVSVKSVANVILYATLAVAGAVTEFLPNTATINEDITIG
APSLVCDPSKAEIGLAPIAAVAKAVGVELTIGDGGSGLTEADRKSSATVTPRENDI
IAAILYSGTGTGRSGKAMLTNDNLANSISLVGKRFPTDQVLIHALPLHYHGHLEAV
TVAITLFSRASMTFLKLPDPLITIKLMARSLVMGVPETFTRLLONAALSRRTYRMRI
FISGSAPILAEIHEWMSARTGHAVALERYGMTETNNMTSNPYGDERVPGCAVPTLGVG
LVTPDPTKELPREIRIGMIEYVKGPNVEFGYWRMPEKTAERPDGQFTIDGLGKID
KCYVHLIGRGKDLVTSIGGFNNVYKPEIESIDAMPGVESAIVYGHADGEGVFTVL
CKKGADEVSASVLAALDGRLLAFKKMKRFVYVDELPRNTMGKVVQNNVLLDPTKKDIYAL
K"

```

[illegible]

RESULT	18				
AX081571					
LOCUS					
DEFINITION	Sequence	924 bp	DNA	linear	PAT 27-FEB-2001
ACCESSION	AX081571				
VERSION	AX081571.1				
KEYWORDS	GI:13170390				
SOURCE					
ORGANISM	Moraxella catarrhalis. Moraxella catarrhalis Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;				
REFERENCE					
AUTHORS	1 (bases 1 to 924) Bethel,F.X., Dalemans,W.L., Denoel,P., Deguesne,G.S., Feron,C.S., Lobeet,L.S., Poolman,J.S., Thilly,G.S., Linnard,J.S. and Voet,P.S.				
TITLE	Genetically engineered Bieb vaccine				
JOURNAL	Patent: WO 0109350 A 76 08-FEB-2001; (BE)				
SMITHLINE	BEECHAM BIOLOGICALS S.A.				
FEATURES	Location/Qualifiers				
source	1..924				

[illegible]


```

/codon_start=1
/transl_table=11
/product="DNA gyrase beta-subunit"
/protein_id="A049900.1"
/db_xref="GI:1899004"
/translation="MSDNVEYGAQIOVLEGLAARPRAMTIGSTDSRGLHLYEV
VNSIDELAGHCDIAEVALHEDGVSATNDGRGIPVGHEDYDRALEIVMTVAJAG
GRFNDKSYOVSGLHGVTVYNAHSSSELEVYKHDGAWTHREFYGEPRVFEPRV
TFLREGIREFVEYLYNEKTALHDVIVYDSESEGLEVEIAQADELOGSITHARN
INTRREGTHLTFKFTALTRYNDYANDSHMDLDDNLGDEVRGLTAVISVKNPD
POEGOTKTKLGSVEVGIYESVTHOOLGTFEENEDTATIIISKAVEARAKAKO
AELTRKRSALSTSLPGKLDACSGSPSESELFIEGDSAGSACOGDRKFOALP
LKGKTIINVEKRLDRILENDEIRALITAGGVGDEFEIEKARYORLIMTADVGA
HRTTLITLTKRMRPLIENGYYAQAAPPLYRVRVRYGNTYDAMDREDRITIEECNG
NPTVOVRFKGLGEMFDPQIMDTYMPENRYLKRITVEDAAADRFNIIIMGDAVPRK
QFKDHANDAEWDI"
/join(7680..9983,1..893)
/organism="Shuttle vector pMD20"
/db_xref="taxon:56961"
/plasmid="Bluescript pBS+"

```

```

BASE COUNT      2058 a      2992 c      2880 g      2053 t
ORIGIN

```

```

Query Match      4.6%; Score 41.4; DB 12; Length 9983;
Best Local Similarity 47.8%; Pred. No. 26;
Matches 120; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```

```

QY 379 cacatcgagctacagatttggcgagcgtacacagcagcagcttcgctccacgtg 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5444 CACGACCTCGTCTAGCAAGTGTGTCGACAACTCCATCGACGAGGGCGGCGGCGCATGC 5503
QY 439 accgcacatgtacacgagccgcgaataatcagaatagacaaatcagcagcgagcg 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5504 GACGCGATCGAGTGTGCGCTCCACGAGAGCGGCTCTGTACCGCTACCGACGAGCGCGC 5563
QY 499 gtgcgcgcaagcaagcaaacacgcgcgcacgcgcacagcagcagcagcagcagcag 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5564 GGGATTCGGGTAGTACCGCAGACAGTACGACCGCGCGTGGAGTGTATCATGACGC 5623
QY 559 gccctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5624 GTCTCCACGCGCGCGGAGTGTGACAACTCTTACCAAGTCTCGGCGCGCTCCAC 5683
QY 619 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5684 GCGGTGCGCGT 5694

```

```

RESULT 23
HLFINSVECB      11291 bp      DNA      linear      BCT 07-MAR-1995
LOCUS      Haloferax volcanii gene.
DEFINITION      Haloferax volcanii.
ACCESSION      L26978
VERSION      L26978.1
KEYWORDS      GI:460656
SOURCE      Haloferax volcanii (strain Aa2.2) DNA.
ORGANISM      Haloferax volcanii;
                Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
                Halobacteriaceae; Haloferax.
REFERENCE      1 (bases 1 to 11291)
AUTHORS      Holmes,M., Pfeiffer,F. and Dyal-Smith,M.
TITLE      Improved shuttle vectors for Haloferax volcanii including a
                dual-resistance plasmid
JOURNAL      Gene 146 (1), 117-121 (1994)
MEDLINE      94341559
COMMENT      L26978 shuttle vector pMLH3
                bases 187..4290 Haloferax Aa2.2 PHK2 replicon
                bases 187..4290 Haloferax Aa2.2 novobiocin resistance determinant
                6973..8978 Haloferax volcanii mevinolin resistance determinant
                8978..187 Escherichia coli vector pBS+.
FEATURES      Location/Qualifiers

```

```

source      1..11291
            /organism="Haloferax volcanii"
            /strain="Aa2.2"
            /db_xref="taxon:2246"
            rep_origin      1..4290
            /note="Haloferax Aa2.2 PHK2 replicon"
            rep_origin      9029..11290
            /note="E.coli vector pBS+"
BASE COUNT      2155 a      3502 c      3429 g      2205 t
ORIGIN

```

```

Query Match      4.6%; Score 41.4; DB 1; Length 11291;
Best Local Similarity 47.8%; Pred. No. 26;
Matches 120; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```

```

QY 379 cacatcgagctacagatttggcgagcgtacacagcagcagcagcttcgctccacgtg 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4737 CACGACCTCGTCTAGCAAGTGTGTCGACAACTCCATCGACGAGGGCGGCGGCGCATGC 4796
QY 439 accgcacatgtacacgagccgcgaataatcagaatagacaaatcagcagcgagcg 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4797 GACGCGATCGAGTGTGCGCTCCACGAGAGCGGCTGTGACGCGTCCACGACAGCGCGC 4856
QY 499 gtgcgcgcaagcaagcaaacacgcgcgcacgcgcacagcagcagcagcagcagcag 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4857 GGGATTCGGGTAGTACCGCAGACAGTACGACCGCGCGTGGAGTGTATCATGACGC 4916
QY 559 gccctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4917 GTCTCCACGCGCGGAGTGTGACAACTCTTACCAAGTCTCGGCGCGCTCCAC 4976
QY 619 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4977 GCGGTGCGCGT 4987

```

```

RESULT 24
AP003906
LOCUS      AP003906      112428 bp      DNA      linear      HTG 12-JUL-2001
DEFINITION      Oryza sativa chromosome 8 clone OJ1191_D12, *** SEQUENCING IN
                PROGRESS ***, in ordered pieces.
ACCESSION      AP003906
VERSION      AP003906.1
KEYWORDS      HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (cultivar:Nipponbare) DNA, clone:OJ1191_D12.
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                Eriaristidae; Oryzae; Oryza.
                1 (bases 1 to 112428)
                Sasaki,T., Matsumoto,T. and Yamamoto,K.
                Oryza sativa nipponbare(GAS) genomic DNA, chromosome 8, BAC
                clone:OJ1191_D12
                Published only in Database (2001) In press
                2 (bases 1 to 112428)
                Sasaki,T., Matsumoto,T. and Yamamoto,K.
                Direct Submission
                Submitted (11-JUL-2001) Takuji Sasaki, National Institute of
                Agrobiological Resources, Rice Genome Research Program; Kamondai
                2-1-2, Tsukuba, Ibaraki 305-8602, Japan
                (E-mail:tsasaki@abrc.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
                Tel:81-298-38-7441, Fax:81-298-38-7468)
                The nucleotide sequence of this BAC clone was generated by
                combining Monsanto and RGP-Japan sequencing data.
                NOTE: It currently consists of 1 contigs. Gaps between the contigs
                are represented as runs of N. The order of the pieces is believed
                to be correct as given, however the sizes of the gaps between them
                are based on estimates that have provided by the submitter. This
                sequence will be replaced by the finished sequence as soon as it is
                available and the accession number will be preserved.
                * NOTE: This is a 'working draft' sequence.
                * This sequence will be replaced

```

```

REFERENCE      1 (bases 1 to 112428)
AUTHORS      Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Direct Submission
JOURNAL
COMMENT

```


* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

Location/Qualifiers
1..112428

/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="8"
/clone="OJ1191.D12"

BASE COUNT 32676 a 23073 c 22916 g 33569 t 194 others
ORIGIN

Query Match

Best Local Similarity 46.6%; Score 41.4; DB 2; Length 112428;
Matches 132; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

Oy 343 ttggacaaggcggaaggctgttcaatcaagccgcaatcgagctacgattggc 402

Db 108495 TTGCAATGAGCGAGGTGAGCGCGCGCTGCACCGCTCATCCGCTCGCATGAGCGCTC 108554

Oy 403 ggaagctacatcaagcagcagcttccttcacacgtgacccgcatgataaagccgcgaa 462

Db 108555 GACCGCGCGTGCAGCCCTTACACGCTTGTCTACAGCGGCTTCAGAGCGCGCACG 108614

Oy 463 atcaagcgatagaacaaatcatgcaagcgagcgaggtgcgcaaggaacacgcg 522

Db 108615 GCGGTGCGCCACGACGACGCGCGCGGTGCGCGCGGCGGACGAGCGCGCGCC 108674

Oy 523 cccacggcgatagaaggggttaacaaatcatcaagcgctgcgagcgaggaacacgcg 582

Db 108675 CTCGCGCGCGTCTGCGGACCGTGCAGCGCGCGGCGGACGAGCGCGCGCTACAC 108734

Oy 583 atcatctgccccgacacgcttccttcgcgaggaagcgagcg 625

Db 108735 CGCATCGTCTCCAGGCTCTCGAGCGCGACCGCGCGCGCGCG 108777

Oy 583 atcatctgccccgacacgcttccttcgcgaggaagcgagcg 625

Db 108735 CGCATCGTCTCCAGGCTCTCGAGCGCGACCGCGCGCGCGCG 108777

RESULT 25 138881 bp DNA linear HTG 27-DEC-2001

AP004589 Oryza sativa chromosome 8 clone P0556A11, *** SEQUENCING IN

LOCUS AP004589

DEFINITION Oryza sativa (cultivar: Nipponbare) DNA, clone: P0556A11.

ACCESSION AP004589.1 GI:18146736

VERSION HTG; HTGS_PHASE2.

KEYWORDS Oryza sativa

SOURCE Oryza sativa

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 138881)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

REFERENCE Submitted (26-DEC-2001) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kamondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,

Tel: 81-298-38-7441, Fax: 81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces is believed

to be correct as given, however the sizes of the gaps between them

are based on estimates that have provided by the submitter. This

sequence will be replaced by the finished sequence as soon as it is

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

Location/Qualifiers
1..138881

/organism="Oryza sativa"
/cultivar="Nipponbare"

BASE COUNT 40689 a 28994 c 28928 g 40069 t 201 others
ORIGIN

Query Match

Best Local Similarity 46.6%; Score 41.4; DB 2; Length 138881;
Matches 132; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

Oy 343 ttggacaaggcggaaggctgttcaatcaagccgcaatcgagctacgattggc 402

Db 530 TTGCAATGAGCGAGGTGAGCGCGCGCTGCACCGCTCATCCGCTCGCATGAGCGCTC 471

Oy 403 ggaagctacatcaagcagcagcttccttcacacgtgacccgcatgataaagccgcgaa 462

Db 470 GACCGCGCGTGCAGCCCTTACACGCTTGTCTACAGCGGCTTCAGAGCGCGCACG 411

Oy 463 atcaagcgatagaacaaatcatgcaagcgagcgaggtgcgcaaggaacacgcg 522

Db 410 GCGGTGCGCCACGACGACGCGCGGTGCGCGCGGCGGACGAGCGCGCGCTACAC 351

Oy 523 cccacggcgatagaaggggttaacaaatcatcaagcgctgcgagcgaggaacacgcg 582

Db 350 CTCGCGCGCGTCTGCGGACCGTGCAGCGCGCGGCGGACGAGCGCGCGCTACAC 291

Oy 583 atcatctgccccgacacgcttccttcgcgaggaagcgagcg 625

Db 290 CGCATCGTCTCCAGGCTCTCGAGCGCGACCGCGCGCGCGCG 248

Oy 583 atcatctgccccgacacgcttccttcgcgaggaagcgagcg 625

Db 290 CGCATCGTCTCCAGGCTCTCGAGCGCGACCGCGCGCGCGCG 248

RESULT 26 204050 bp DNA linear BCT 07-DEC-2001

AL646073 Ralstonia solanacearum GM11000 chromosome, complete sequence;

segment 17/19.

ACCESSION AL646073.1 GI:17429991

VERSION AL646073.1 GI:17429991

KEYWORDS Ralstonia solanacearum.

SOURCE Ralstonia solanacearum.

ORGANISM Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

Ralstonia.

1 (bases 1 to 204050)

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S.,

Arlot, M., Billault, A., Brottier, P., Camus, D. C., Cattolico, L.,

Chandler, M., Choise, N., Claudel, R., Robert, C., Cunac, S., Demange, N.,

Gaspin, C., Lavi, M., Moisan, A., Robert, C., Saurin, M., Schlex, T.,

Signier, P., Thebaud, P., Whalen, M., Winkler, P., Levy, M.,

Weissenbach, J., and Boucher, C. A.

Genome sequence of the plant pathogen Ralstonia solanacearum

Unpublished

2 (bases 1 to 204050)

Boucher, C. A.

REFERENCE Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston

Cremieux, CP5706, 91057 Evry Cedex, France. Laboratoire de Biologie

Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,

Bp27, 31336 Castanet-Tolosan Cedex, France, Fondation Jean

Dausser-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, LMG CNRS

118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA

URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biométrie et Intelligence Artificielle INRA, Bp27,
F31326 Castanet-Tolosan Cedex, Laboratoire de Génétique Cellulaire
INRA, Bp27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html.

FEATURES

Location/Qualifiers
1..204050

/organism="Ralstonia solanacearum"
/strain="GM11000"
/db_xref="taxon:305"

gene
complement(115. .1278)
/gene="dgt"
/note="RSC2968; RS01329"
complement(115. .1278)
/gene="dgt"
/EC.number="3.1.5.1"
/function="small molecule metabolism; central intermediary metabolism; nucleotide hydrolysis"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE DEOXYGUANOSINE TRIPHOSPHATE
TRIPHOSPHOHYDROLASE PROTEIN"
/protein_id="CAD16677.1"
/db_xref="GI:1742992"
/translation="MNDRLDLPAPDDECHLAPYAANAQSGRVHAPSPSTREF
QDRDRITICTAFRLRYKTVFVNHGDLFRRLTHSEVAQIANSIRSLRNEDL
VEAISLAHDGHTPEAGQDALNDCMKPYGFEHNLQSLVDRLEERYGGDGLNT
TFETREGLIKHCRRNNAATLGLDRERLEQPSLEAQLANLADYVANNHIDDLAI
SGLITTELEGVPLMARHREAEAFPGVGRMI NETIRITINALIDLAGTAAI
TGAAPSIDVVRAPPLVAREPMDREARLKRFLDNLRYHYLVMRMAKARITIED
LRFVETDPLRLPQIQARPAQPKRIAHYTAGMTDRAIKERIRIFAEVYL"
complement(1298. .2404)
/gene="arob"
/note="RSC2969; RS01328"
complement(1298. .2404)
/gene="arob"
/EC.number="4.6.1.3"
/function="small molecule metabolism; amino acid biosynthesis; chorismate"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE 3-DEHYDROQUINATE SYNTHASE PROTEIN"
/protein_id="CAD16678.1"
/db_xref="GI:1742993"
/translation="MITVDYDGERAPTHIGTILSQAELEFAPHRGRAVIVTNET
VAPLYARVEAIIKSGTIVDMVYLPDGESEFKWETINRIFDMLILASGADRTTVAL
GGGVIGMTGFPAASVYRGVFIQVPTTLISQVDSVSGKTGINHPLGKMICAFHQP
QAVLADIDTLRTLPRELAAGMAEVIKGAITADYFAWIERHIAGINACDADLMGA
VRGSVOIKAAVADDERESGLRALINFGHFGAIEAGLGYEMLGEAVGCCMAMA
DLSHRLEFIDIDTRNVTALTALRAANLPVAPPDGVARFIDLMRVDKKAEAGEIKFVLL
RKIGSAFTVTPDRLATLOHNLVRPTTAPVA"
complement(2401. .2973)
/gene="arok"
/note="RSC2970; RS01327"
complement(2401. .2973)
/gene="arok"
/EC.number="2.7.1.71"
/function="small molecule metabolism; amino acid biosynthesis; chorismate"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE SHIKIMATE KINASE I PROTEIN"
/protein_id="CAD16679.1"
/db_xref="GI:1742994"
/translation="MSVSNVFEVGLMGAGKTYTGRAAARRLDLPFDSDHEIARCV

gene
complement(3114. .5258)
/gene="p110"
/note="RSC2971; RS01326"
complement(3114. .5258)
/gene="p110"
/function="structural elements; cell exterior; surface structures"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE FIMBRIAL TYPE-4 ASSEMBLY SIGNAL PEPTIDE PROTEIN"
/protein_id="CAD16680.1"
/db_xref="GI:1742995"
/translation="WTMRVKGIVQAVRVARGGAVVAMLSIMLMGGOALAOAITA
PASNAVERVEGASVESTVAVITLKDTPAQKPVFEFTQOPARAIIDPFGAPRQGRAN
YQGGKILRAASVQIGRTRVVLDLRAQSOYKSEIGNOVILTLGAAPTASATPVPT
FAAPAP7AGVRSVSRINIDFRGSELARVVLDLSTNSAINTIAQOQNVVDPGAT
LPQSLRRFEDVSDGTPVQARATDNGALVLEPGMNOQSYQDQFVVEVRT
KEDPKLISGPGYGERMSLNFONIDISLQVADPFLNLTIVSDVGTGLSLRLD
VPMQDALDIVLDSKGLASRRNGNVLVAPRRELATKRAELSEGOQVTEPLRSQVE
RLNFORADVDVRLMLGTGTGGGAASRIILSKSGISLSPARTQQLVSDIPSKLEEV
QAFILKIDIPVQVMTIARIVADDTESRNIGAKLGEPAKTNAGYCNITSNVSPPT
KDATWNSAAVSLPANGINGVNAASVAVSLPNAAGRLALELSALDEAGKLTSSP
RYVTDNLIKALIEQTEVYQOATSSGATSVFKRANKLEVTPKIPDGVFLVDVY
NKDSIGTQTNGFAINTKHVOTQVIVENGGVVIGIYQNERFDVKKVPLLDGIPVL
GNLFTAKTNDREPLVLELPRVLSQLSLK"
complement(5255. .5800)
/gene="p1lp"
/note="RSC2972; RS01325"
complement(5255. .5800)
/gene="p1lp"
/function="structural elements; cell exterior; surface structures"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE FIMBRIAL TYPE-4 ASSEMBLY LIPOPROTEIN"
/protein_id="CAD16681.1"
/db_xref="GI:1742996"
/translation="MKGOVMTKROVTRLIHLSPLCLAVIYAACASADELEQWMOQA
RSTVSRFLAPLPEPKVAPREVYTKQTPFSVOGAEISRAQNSPPGRRPRLD
VPLAEFLKLTGTRKNGSESAVLSVGDQKHVVGQYIGONYGRIYIDQDLTLRELV
RGTIAEMKEMTKLVQSGAC"
complement(5797. .6462)
/gene="p110"
/note="RSC2973; RS01324"
complement(5797. .6462)
/gene="p110"
/function="structural elements; cell exterior; surface structures"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=1
/evidence=not_experimental
/product="PROBABLE FIMBRIAL TYPE-4 ASSEMBLY MEMBRANE TRANSMEMBRANE PROTEIN"

Query Match Best Local Similarity 4.6%; Score 41.2; DB 1; Length 204050; Matches 112; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

gene
CDS

protein_id="CAD16682.1"
db_xref="GI:17429997"
translation="MTMLSTELSLSEELYVGFRCGLNLEPFWOLAPILFMIIAIVLVVGWQFTWGSKLDERAKRTDEQLKSAVOSKIQAOVNLAEALKOKI EYERRPAKIELQENRTEMALLADVNHAGIARGLTPEASAVVKPYVAELIPSVKYNGRHNDALFADVVALSRIVYLIRNIALTGNKGSGMWABRAEAYRALDPDQAAQRKAAYVKDAK"

/gene="plin"
/note="RSC2974; RS01323"
complement(6459..7091)
/gene="pin"
function="structural elements; cell exterior; surface structures"
/note="product confidence : probable

522 gccccacggcatcacacaagggttcatacaaatcatacgaagcccgcgcgagagcaac 581
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 188183 GCTGTGCGGGCCGCCAGGGTTTGAGGCCCTTCAGAAGGCCTTTCGAACCGCGCGGCC 188124

Qy 582 cctatcattgccgcgaccacgcttcccctttccgcagaagaagcgcgagtgttggcgatatt 641

Db 188123 GACCCTCACATGATCAACGCGTTGCCACGCGCGAGAGGTCGCCAACCTGTGTGCTCATGT 188064

Qy 642 ttccgcgaacctgatatcacccaatgcacatgcactgccccgagcaaaattggcaacggttaaggcgt 701

Db 188063 GGCCAGCCCCCGCTGTGTGTGCGGCCACCAACAAGGCCGCCGCTGCGGGTGTGATGCGCGGCTGT 188004

Qy 702 gaaaacctgtttctctgctgcgaagcgctgccccgagcgagaagcgcttg 751
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

Db 188003 GAAGTCACGCGTCTGAGCGCCGCCGTCGCCGCGCGGTCGTCATGCTCCG 187954

RESULT 27

LOCUS AB0383378

DEFINITION Streptomyces lividans cpbl gene for chitinase promoter binding protein 1, complete cds.

ACCESSION AB0383378

VERSION AB0383378.1 GI:9857261

KEYWORDS chitinase promoter binding protein 1.
Streptomyces lividans (Strain:66 TK24) DNA.
Streptomyces lividans

SOURCE Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

ORGANISM Fujii,T. and Miyashita,K.

REFERENCE 1 (bases 1 to 1760)

AUTHORS Fujii,T. and Miyashita,K.

TITLE A DNA binding protein that bind to chitinase promoter region In

JOURNAL Streptomyces lividans Published Only in Database (2000) In press

REFERENCE 2 (bases 1 to 1760)

AUTHORS Fujii,T. and Miyashita,K.

TITLE Direct Submission

JOURNAL Submitted (15-FEB-2000) Takeshi Fujii, National Institute of Agro-Environmental Sciences, Laboratory of Soil General Microbiology, 3-1-1 Kan-nondai, Tsukuba, Ibaraki 305-8604, Japan (E-mail:takeshi@naes.afric.go.jp, Tel:81-298-38-8256, Fax:81-298-38-8199)

FEATURES

CDS source Location/Qualifiers

1..1760 /organism="Streptomyces lividans"
/strain="66 TK24"
/db_xref="taxon:1916"
BBS 341..345
gene 356..1213
/gene="cpbl"
356..1213
/gene="cpbl"
/codon_start=1

[illegible]

gene
CDS
/db_xref="taxon:69394"
complement(114..1757)
/gene="CC0685"
complement(114..1757)
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="chaperonin, 60 kDa"
/protein_id="AAK22670.1"
/db_xref="GI:13421906"
/translation="MAAKDVFSDDAKDKMLRGVNIILANAQVTLGPKGRNVIEKSG
GAPRTYDGVSAKEIDADLFENLGAQMTREVAASKTNDKAGDATTAVLAQAIQVE
GKSAAGMNPMDLRGIDRAVAIAIEDITSKKTNTAEIAQVGTISANGKEVGE
MIKAMKRVNEGIVTEAEKTAETELDVEGQPDGYSPEITNADMEVQLEEP
LILFEKLSISQDPLPVLAVVOSGPRILLIADDEGSELTALVNNKLRGILRAVAV
KARGSGRRKAMEDDILILGAQVSDSDITKIENNSLEMLAKKAVSTIKDPTIYD
GVGSKADIEARIQKIQIEDTTSYDIEKLEKELAKAGVAVIRNGVSTVEVEK
KDRYDALNMTAADEGIVGGGTALAKSKALAGVGDNDQTAGIALVRALQAP
IRQIAENAGVSGSVGKILENDSAFGNQTEOYVDVLDVDPKAVRTALQNA
ASVAGLITTEAEIVAEPAKKGAGPAGGMPGGMDF"
complement(1815..2105)
/gene="CC0686"
complement(1815..2105)
/gene="CC0686"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="chaperonin, 10 kDa"
/protein_id="AAK22671.1"
/db_xref="GI:13421907"
/translation="MKRPLDRLVLRVEEETKGGIIPDTAKERPOEGEVAVAG
PGANDKGVVADLVKAGDRILFKMGSTGVKVDGODLLIMKESDVLGYVEA"
2280..2474
/gene="CC0687"
2280..2474
/gene="CC0687"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK22672.1"
/db_xref="GI:13421908"
/translation="YMDHGVYVNDLQASRAFGATLAAGLDNGAGVRPHYHY
GAFVLDPDGRNVEAVCHAPA"
2493..2921
/gene="CC0688"
2493..2921
/gene="CC0688"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK22673.1"
/db_xref="GI:13421909"
/translation="WSSYVLSLLAVIACVIGTGLGMIARPATLIEAAGSNDTP
KPLFASGRALGVILASHGVALLYLOPRIGAMALALALAMLGAAGRAVSLAD
GDAGRNYGAMLFNALIGVTLFPMFQIGKIVLGRPLGLA"
complement(3122..3784)
/gene="CC0689"
complement(3122..3784)
/gene="CC0689"
/note="similar to PID:769899; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="methylated DNA--protein-cysteine
methyltransferase"
/protein_id="AAK22674.1"
/db_xref="GI:13421910"
/translation="MRGCGSPFVKGVPVPLIRSPSPSPARGNRYNSRAR
LGDHMRPRANHLFLDRAPSPGLLITPTEGTALVYDFEPRRLARLRYHSDR
VLETPAPAIKAFAKRFEGDLALAVSLMAITNGTFQDAVMTALQIIPAGOTIYS

gene
CDS
complement(3835..5142)
/gene="CC0690"
complement(3835..5142)
/gene="CC0690"
/note="identified by match to PFAM protein family HMM
PF00441"
/codon_start=1
/transl_table=11
/product="acyl-CoA dehydrogenase family protein"
/protein_id="AAK22675.1"
/db_xref="GI:13421911"
/translation="MDPDISPKORAFILDRVTAFMDEHYPAIPAEEMVLAERKK
VVOVEDLKKAKAAGLWNPFPSCQTHDDTFQEGVQVLTMLYSLAEQLKVG
FASEVNSAPDPTGMVLEMRGTLAGCKEMLRPLMNGEISAFLEMPAVASDYN
TETRIERGDHYVINGRKWSSGVGDPCYKAIYWKDPNARHQAQSVLPPLS
PEIETVRMLPVTGYTDARHGAHYLVKYSVSDSHVBERVADARDIEMCLLCKKA
HMKRTIGTAIEVALEKRCORLAKRKATGKYSVSDSHVBERVADARDIEMCLLCKKA
DMMDRAGKSALEIATMTKVAAPRLAKI1DDAIOAHGGGVTTDFGLAKAYAGIRTL
RLADPDEVHQRITARMYGRYDLMYKQAKKEKALHVPGR"
5269..5895
/gene="CC0691"
5269..5895
/gene="CC0691"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK22676.1"
/db_xref="GI:13421912"
/translation="WPIITVDGPMNLNTADILGRYVLRPEVYVDHDEKAAIDCD
PASWEIMSGEGEGEDFVNGALQETDRGERIGAIIRLVGKVGSTSLNTRRH
GGLGATFLENPARGSPVNPESKRLMGHAFDKGALIRVLDVDRNARSQAIOKL
GATKEGVLRNHKVTWGTGVRDTAVPSITDYPALRERLETSER"
5968..7131
/gene="CC0692"
5968..7131
/gene="CC0692"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK22677.1"
/db_xref="GI:13421913"
/translation="MAGMACLRICIEHGARSVTLGDTVPANIKVASYIDVNRAPR
RHGRVPGAYVLIDPSAETLNGEELROMAAELOVILFCTGQODICMLTFEGDEOVL
KPTTLPEEELAMRGAEPPPEGRTRIVSODGRIELAPMGSTPVAEAEQDEOVL
RLGMRGVSTPSORPEAAOYOSFLKPGAEPSQSDPMDRLOALDRALVALEDPRL
KAMVGRLMTVARGLRDTCCRRFAITPSVREIRGAVYDAPRDPSPSALGPLRDLY
QPYFARVGLCVNDPQFYVALDVLHCVFQLPEDERKARVLAHRLSARGQYVAR
SLRMGLAGVKTPTRELELARKMGHMYKGPMSIALFATPVSDVVPELPLRT"
7154..8530
/gene="CC0693"
7154..8530
/gene="CC0693"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK22678.1"
/db_xref="GI:13421914"
/translation="MAAGCTKSITATLNMGRVNAHAAGMTAIAICRILITADADRAG
DMPKRAOGVADAFKAKWRKPRKGAVALAFVLAADNATELPTVEQGLAVELEV
LPPNPRTGIALMTEFGEDETAVLHFSALSOKELASIMAGEGSGPGEVHVVTADAVK
PMPVTRERKVVAAEKARAAEAKRPAAPAAAPAAVAPAPSPAKSAPAPPTAP
VAAKYVTRPRADITGKMGVYDLSQAFPGSVYAMADIVPPAPRDLAKLRDILVAT
DVQALKTASRGEIOASGFNWLITPASQESTKALSTPLDAPRLTASITYGAREA
SIGMDQMRQALSGAFAMDLRTIDTPTPVVGLPELLESTVYLVISGETEDRKQIL
KEVERKSAYAAKGVQCVANASAVELEACRAGVAVVQGVAVIVDMDTPPAAADGAP
LALAKSAA"
8606..9373
/gene="CC0694"

	Accession	E32986			
	Version	E32986_1	GI:13022340		
	Keywords	JP 2000060568-A/1.			
	Source	Vigna angularis			
	Organism	Vigna angularis			
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae Vigna.			
REFERENCE	TITLE	1.(bases 1 to 10732)			
AUTHORS	Koiuchi,M.T.K.K. and Sato,D.S. Gene encoding cellulose synthelizer				
JOURNAL	Patent: JP 2000060568-A 1 29-FEB-2000;				
COMMENT	KOICHI MIZUNO,MITSUI GIYOUSAI SHOKUBUTSU BIO KENKYUSHO OS Vigna angularis PN JP 2000060568-A/1 PD 29-FEB-2000 PF 26-AUG-1998 JP 1998239998				
	PI KOICHI MIZUNO,TOMOHICO KATO,SHIGERU SATO,DAISUKE SHIBATA PC C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10,C12P21/02// PC (Cl2N15/09,(Cl2R1:91),(Cl2N5/10,C12R1:91),(Cl2P21/02,C12R1:91), PC C12N5/00, PC C12N5/00,(Cl2N15/00,C12R1:91),(Cl2N5/00,C12R1:91) CC FH key location/Qualifiers FT CDS (1). .(3375).				
FEATURES	source	location/Qualifiers 1..10732 /organism="Vigna angularis" /db_xref="taxon:3914"			
BASE COUNT		3149 a 1212 c 2074 g 2046 t 2251 others			
ORIGIN					
Query Match	Best Local Similarity	4.4%; Score 39.6; DB 6; Length 10732;			
	Matches	79; Conservative 246; Mismatches 274; Indels 1; Gaps 1,			
Oy	147	ttaaagaagccgcgcgatcgtgccaaatttggcgaaggcttgtaacccca 206			
Db	9474	TTTHRKGTHTRAASRTSGSGSRGNGASHSYSGYHSRRVVGHSSRSASNVAARSRGSHR 9533			
Oy	207	cacgcgaagcgytcacaagccgttttggcgaagcaagcgaagtgcgttggactggccc 266			
Db	9534	HAAHSRASAGYSNASASYSTRHRCTCSMAOSTTRARSAASGRASHAHGSKGYSTYS 9593			
Oy	267	cgcgttttcaaataaacccgaagacatcgaataatgtlccaagcgcgatcacgcgtggga 326			
Db	9594	YSGRRCYSTSYSRVAVGAASRASAMNTMASHVVGMTGYSNAVAVASRVAGYSMAATHRG 9653			
Oy	327	aacacgtgcagaagcgtttggaccaaggggaagggttgtctgtlccaagccgcacatcgg 386			
Db	9654	NNSRRASAASAVAATHTIRGASNSKRKTTHGGYTHRHVSGRTHRGGGYSGGRAVARSHS 9713			
Oy	387	cagctaagatttggcgcagcatcatacgaacgaagcttcgtlc-caactgaagccga 445			
Db	9714	SRGWTHTRTNRGTHTTTHRGVAANAAMTGMRMTPSAMSSHRSARRTSRYKSANTHCYSR 9773			
Oy	446	tgtaacaagccgcgaaaatacaaacgcataagaacaaatatatgcgaagcgggcagylgcgq 505			
Db	9774	ASHTSGTTHRAYARTITROMASVAAYASTHKGANSNHNVATHRHVSQGYRAMASHSAYS 9833			
Oy	506	gcaaaagccaacacccgcgcaccaacgatalaagaaggtttaaacaatcataaagccccgcg 565			
Db	9834	VAAAARGHSVAAANASHASHASHSMASRRANGYSRVAAHASMTHRGANSNRMGCHSHSC 9893			
Oy	566	gcgcggcgccggagcaacatcatctctgcccgacaacgcgtcccltccgaagaaagcgccq 625			
Db	9894	YSSKCYSRRASSRRFVNVAASNNRRAAGYGGSNMMAAASRYEARCGRRGNCAASERYSRG 9953			
Oy	626	gcgcgttgcgcgaatttttccgcaaaccttatcacacatgcacctgacgttcggcgaatctg 685			
Db	9954	YXSCVAGAABRAAVAASTAASRSSRMSRKNYSYGTARKGASHRTYTRASNGCYMTPTH 10013			

	QY	686	cacagctcaaaaggcgtgaaccgccgttttcttgtcgagaacctgcgacggacaag	745
			: : : : : : : : : : : : : : : : : :	
	Df	10014	Rcyrrmtyvaangvahnrtistrythsrgksritysvaghhssactytryrassrgytsthrs	10073
RESULT	36			
AEO04441				
LOCUS		12213 bp	DNA linear	BCT 30-AUG-2000
DEFINITION		Pseudomonas aeruginosa PA01, section 2 of 529 of the complete genome.		
ACCESSION		AEO04441 AEO04091		
VERSION		AEO04441.1 GI:3945828		
KEYWORDS				
SOURCE		Pseudomonas aeruginosa. Bacteria; Proteobacteria; gamma subdivision; Pseudomnadaaceae;		
ORGANISM				
REFERENCE				
AUTHORS		1 (bases 1 to 12213)		
		Stover,C.K., Pham,X.Q., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Huftagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y., Brody,L.R., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.		
TITLE		Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen		
JOURNAL		Nature 406 (6799), 959-964 (2000)		
MEDLINE		20437337		
AUTHORS		2 (bases 1 to 12213) Stover,C.K., Pham,X.-O.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S.L., Huftagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y., Polger,K.R., Kas,A., Larbig,K., Dim,R.M., Smith,K.A., Spencer,D.H., Mong,G.K.-S., Wu,Z., Paulsen,I.T., Keller,J., Sater,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.		
TITLE		Submitted (16-MAY -2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
JOURNAL		location/Qualifiers		
FEATURES				
source		1..12213		
		/organism="Pseudomonas aeruginosa"		
		/strain="PA01"		
		/db_xref="taxon:287"		
gene		83..634		
		/gene="tag"		
		/note="PA0010"		
CDS		83..634		
		/gene="tag"		
		/codon_start=1		
		/transl_table=11		
		/product="DNA-3-methyladenine glycosidase I"		
		/protein_id="AAG03400.1"		
		/db_xref="GI:3945829"		
		/translation="MPRCWCANDPLIYMAYHDEMGCVPORDDALFEILLLEGFOAGL SMITVTKREKREVLYFGFDVCVRADMSAEIDELMDGIIIRNRKLAAARONAMW LELDPPAGFLMSPVGOPKIHFPAAGRAEVPATPEAKMSKALLRRAGENFVGPTICYA FWASQSWVDHDQDCRYAQLVG"		
		778..1665		
		/gene="PA0011"		
gene		778..1665		
		/gene="PA0011"		
		/codon_start=1		
		/transl_table=11		
		/product="probable 2-OH-lauroyltransferase"		
		/protein_id="AAG03401.1"		
		/db_xref="GI:3945830"		
		/translation="MEKEFGALVALRLRFALLPMRAWGVAGTIGIMLMWKPNRSRE VRIINSCEPFSETEKLIVGOSIMOGIRTLSACAMTWPERKSRTIREVEGMIE VLEAFSLGGDGLVGTSHLGNGEVFNHTYCSTAKITITYRPKLAVIDELLKKROVL GNRVAISPEGDIIVTIYEKKKGCGYGITPADDEPARTAGLFVYLGTALTISKVFOL SGKRARGVFETHAVRIPLDGSSGYVILEAPADMWDYDKDEVSAAMSRELAKTVRAYPSQ		
CDS				

```

gene      YMSMKRFRKRPDGEKKWY"
CDS       1750..2016
          /gene="PA0012"
          /transl_table=1
          /codon_start=1
          /product="hypothetical protein"
          /protein_id="AAG03402.1"
          /db_xref="GI:9945831"
          /translation="MDNORQYRPTLKCIRISHPILFGLMAQTRDLSDTGVYVYKHPD
          LTQLPTGVSIVTGOVDLPIDAPILDMEVYRVDAEAGVGLRFLSEA"
          /gene="PA0013"
          /note="PA0013"
          /codon_start=1
          /transl_table=1
          /product="conserved hypothetical protein"
          /protein_id="AAG03403.1"
          /db_xref="GI:9945832"
          /translation="MSASSPAAMRGTLAFLAHLFIIIASVYLVQLPTLFGWHTTW
          GAFSPFIALADLYLVLLGKGPALVIAVMI PALIASYVSVLFOEAAFGFSL
          ENFTFVARISLSEFLAYVGOILDIQVPLRRSRHMTAPVASTILGLMIDTFEFPF
          VAFWRSDNPMQHWYEIATVDYGVKLSISLLFVPLGMLNGILKMLPGRPOSA"
          /complement(2878..3150)
          /gene="PA0014"
          /complement(2878..3150)
          /gene="PA0014"
          /codon_start=1
          /transl_table=1
          /product="hypothetical protein"
          /protein_id="AAG03404.1"
          /db_xref="GI:9945833"
          /translation="MSQKRSVPSTVEYEPQALIVYHKNASNEMKLESQERLALL
          GLHALERIEVFTGLVPIQVIAHILVTTLGGDALLTLTAHRAFTS"
          /complement(3443..3760)
          /gene="PA0015"
          /complement(3443..3760)
          /gene="PA0015"
          /codon_start=1
          /transl_table=1
          /product="hypothetical protein"
          /protein_id="AAG03405.1"
          /db_xref="GI:9945834"
          /translation="WMTEGLEKMLAGVDNALRFGLGKGYLDAGDAERAHEILQRCV
          EODPKYSAGKMLLKARQAAGDLAARQAQMEQGLATATHGDKQAEKEMTVFLRKIDR
          AKT"
          /complement(3909..5282)
          /gene="trkA"
          /note="PA0016"
          /complement(3909..5282)
          /gene="trkA"
          /codon_start=1
          /transl_table=1
          /product="potassium uptake protein trkA"
          /protein_id="AAG03406.1"
          /db_xref="GI:9945835"
          /translation="MKIIILGAGOVGSLAEHLASEANDITVVDTGDRRLDGLDRD
          IRTVQKASPTVROAGADDADMLVAVTNSDETNNVACOAVATLFTPKIRARVEP
          AYLRITGLFENEAIPVDYLISPEQVYNNYIKRLIEHGALQVIDEAGKQALGIRK
          YGGLVGOELQRLREHMPNDTRVAALIRRRRPITRPGDYIEADDEVFIIAKAHIR
          AVGEMRLDLSYKRIIAGGNGVERIAEIESRYOKIIEKSPLCRNLSDTLDST
          IVLNGASDRDLLENIENIGETDVLALTNDEANIMSLRGAALRGAQVMTLLNNPY
          VDLVQGEIDIAISPOLATIGTLTAHVRGDIIEVHSLRGAALRGAQVMTLLNNPY
          VIGRSINIKILPPTTIGALVREDEVIILAHODTRIIEGSHVLLFVLDKRYIRVERLF
          QAGITFF"
          /complement(5310..6614)
          /gene="PA0017"
          /complement(5310..6614)
          /gene="PA0017"
          /codon_start=1
          /transl_table=1

```

```

          /product="conserved hypothetical protein"
          /protein_id="AAG03407.1"
          /db_xref="GI:9945836"
          /translation="MNPRLAACQALAAVLAGRASLSCALPOLDKYAPPRDGLIOLA
          FGARWQPRLOALAAARLQKPFRAADTDIHALIGLYQLTRIPPHAIGETVCCA
          DKLKQWAKGVLAVALVRAQREGETLLAEVDROPRLGHPWMLKALQAMPEBOLDA
          LCAANNHAPMPTLRVNRHGERDVAIYLAELAEQIKARACDYSHDGIOLAPRVERLP
          GFAEGRYSVQDEAAQALAEQLESAPGORVLDACGAPGTCCHLETQPLAEVAVDL
          EESKIVYRENIQRLGLOSLVAADARVGEEMWDKRPFORILLDACSATGVYIRRHDP
          IKLARKPEDIAALALQGLLDDMLMPTLIEVGVLYATGSAVPEAENSQSIAMFARTP
          GARLDPGPMQKQPRGROLPLQVBEHDSFYAKLIKISAR"
          /complement(6611..7555)
          /gene="fml"
          /note="PA0018"
          /complement(6611..7555)
          /gene="fml"
          /codon_start=1
          /transl_table=1
          /product="methionyl-tRNA formyltransferase"
          /protein_id="AAG03408.1"
          /db_xref="GI:9945837"
          /translation="MSQALRIVFAGPPEFAAEHLKALLDTPHRIIVAYTQPPRPAAG
          OKLMPSAVKSLALEHGLPVMQPOSILRAAEQALALRADLMVVAAGLIPAVLADI
          PRICINSHSLSLPRMGAPIORAYAGDAEAGVYMMQMEAGLDGPMILKYSTPS
          AADVGSILHDLALGPKAVYTAIAGLAGTUBETDODALATYAHKNDDEARLMS
          RPAYELERQVRAFTPMVPVCHTSLADAPLKVAGSLGSGSAPGTILEASRDGLLVACG
          EGALRLTRLOLPBGKPLAFADLYNSRREPAGAVILQ"
          /complement(7610..8116)
          /gene="def"
          /note="PA0019"
          /complement(7610..8116)
          /gene="def"
          /codon_start=1
          /transl_table=1
          /product="polyepitide deformylase"
          /protein_id="AAG03409.1"
          /db_xref="GI:9945838"
          /translation="MATLITLFPDPRRLTIAPVEVVDVADROLIDMPEETYEAG
          IGLAATOVYHVRILVMDLSEDKSEPRFTINPEPELITEDMDQVQDECLSPGYEYV
          DRPOKVRIRKALDRQGNPEEVAEGILAVCTIOHCDHLNGKLFVYLTSLKRDRIKKL
          EKHROQA"
          /8255..9280
          /gene="PA0020"
          /8255..9280
          /gene="PA0020"
          /codon_start=1
          /transl_table=1
          /product="hypothetical protein"
          /protein_id="AAG03410.1"
          /db_xref="GI:9945839"
          /translation="MRKSLVALMLLILASGLAOAVDLREGHDPRTVYRGDTLMDISG
          KFLQPMKPELMHANPOTIOPHILYPPDTSIYVDSQPRVLVNRGESSGTIKLSPK
          IKSTPIAEIAPITPLDKTNSFLANRIVDDERTTSATYIAGANARIVSGTGRITYA
          RGFADGQPAYGIFRQGRVYIDPKTEVLGINADIDGGEVVAIEGVATIALTRTTO
          EYRLDRLEFTEERAIVNSTFMGPESRVEKEIIDVPGVYIQGFVTVTLNRQDRG
          LAEYNVLAIVYKGVTVRVRVGEYSKIPDERAGLIMVRYTKKLSTALVLAASRPLSV
          MBRQNP"
          /9415..10503
          /gene="PA0021"
          /9415..10503
          /gene="PA0021"
          /codon_start=1

```

Query Match 4.4%; Score 39.6; DB 1; Length 12213;
 Best Local Similarity 45.0%; Pred. No. 68; Indels 15; Gaps 3;
 Matches 285; Conservative 0; Mismatches 334;

```

QY 256 gaactgccccgcggttttcaaaaaacggaagacatcgaaacaatgttcaaacgcta 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1012 GAAAGTGCCCTGCGCTGATCTGCGCCGGAAGTCGCTGAGANACATCCCGAGGTC 1071
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 caccgctgggaacacgtgcacgagcgtttggacaagggcgaaagcgtcgttcacacag 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1072 GAAGGCATGAGAGTGTCTGAGAGAACCGCTGGCTTCCGGCGATGCGCTGTCGCATTAC 1131
Oy 376 ccgacatcgagcagctacgattggcgagcagctacacacccagcagcttcgcgttcac 435
Db 1132 AGTACACCTGGGAA-----ACTGGGAATACTCAACCACTTCTATGCTCTTACGCCAAG 1185
Oy 436 ctgacccgcatgtacaagccgcccgaataatcgaatagacaaatcatgacgagcgagc 495
Db 1186 CCGATCATCTTCTATCGTCCGCCCAAGCTGAAGGAGTGAAGGAGTTCCTCAGAGAGCA 1245
Oy 496 agggcgcgagcaaaagcaaaacccgcccagcgagcagcagcagcagcagcagcagc 555
Db 1246 CGCGTCATTTGGGCAATCGCGCACCTTCACCTCCGGAGGCTTCCTCAGTGCATC 1305
Oy 556 aagccctcgcgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 615
Db 1306 AAGGAATGAAAGAAAGGCGGTTGGCTAGGATTCGCCGCCGCCGCCGCCGCCGCTAC 1365
Oy 616 gaagcgagcgcggtgtggcgagcttttcgcaaacctgacatcacatgacagctgagc 675
Db 1366 GCTGGGCTCTTCGTGC-----CCTACCTGGGCAACACTGATGATGACGAGTTCGTC 1419
Oy 676 gcaaatggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 735
Db 1420 CCGCAGTTCCTTACCGCGCAAGCGCGCTGAGTTCCTTCATGCGCGTCCGCCGCC 1479
Oy 736 gacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 795
Db 1480 GATGGTACCGGTTACAGAGGTATCTCTGAAAGCGGCTCCGGCGACATGTACGACAAAGAC 1539
Oy 796 gccacg---atgcgcgcgtgttcaaccgcaatccgaatatgtgacagcgcttccg 852
Db 1540 CTGGAGAGTCTTAGACGCCATGAGCGCGAGTGGCGAAGTATGTACGAGCTATCC 1599
Oy 853 acgagatctgtttatgtatgaaccgctataaa 886
Db 1600 AGCCAGTATGATGTGAGCATGATGAGCGCTTACAGA 1633

RESULT 37
RCNIF
LOCUS 3570 bp DNA linear BCT 13-MAY-1993
DEFINITION R. capsulatus genes nifU, nifS, nifV, nifW and orf6.
ACCESSION X68444
VERSION X68444.1 GI:297915
KEYWORDS nifS gene; nifU gene; nifV gene; nifW gene; nitrogen fixation.
SOURCE Rhodobacter capsulatus.
ORGANISM Rhodobacter capsulatus.
Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
Rhodobacter.
REFERENCE 1 (bases 1 to 3570)
AUTHORS Masepohl,B.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1992) B. Masepohl, Lehrstuhl fuer Genetik, Univ.
Bielefeld, Fakultat fuer Biologie, Postfach 100131,D-4800 Bielefeld
1, FRG
REFERENCE 2 (bases 1 to 3570)
AUTHORS Masepohl,B., Angermuller,S., Henneke,S., Hubner,P.,
TITLE Nucleotide sequence and genetic analysis of the Rhodobacter
capsulatus ORF6-nifU SVW gene region: possible role of NifW in
homocitrate processing
JOURNAL Mol. Gen. genet. 238 (3), 369-382 (1993)
COMMENT 93261420
FEATURES
Related sequences: M26323 & X07567.
location/Qualifiers
source 1..3570
/organism="Rhodobacter capsulatus"
/strain="B10S"
/db_xref="taxon:1061"
96..110
/note="12/-24"
147..467
promoter
CDS

```

```

/Note="ORF6"
/codon_start=1
/transl_table=11
/protein_id="CAA4848.1"
/db_xref="GI:297916"
/db_xref="SWISS-PROT:007184"
/translation="MEITROPPOPEAKIAIEGACGPVAGLRIMVAGCAGLKGKMAL
ELEADPDSDSLFGEVTVLIDPLTREHIEGRTIDFVNSDEAGFTFDPNPKSACGCC
KSEC"
478..885
/gene="nifU (copy1)"
478..885
/gene="nifU (copy1)"
/codon_start=1
/transl_table=11
/protein_id="CAA4848.1"
/db_xref="GI:297917"
/db_xref="SWISS-PROT:007178"
/translation="MRDQDDDTKSPADPPAAARRAAGAPDASALDRPAKLA
QADTPAADAAADDDTTRIRALIDEMRPTFRDGGDIELVREGAKVIVHSGAC
ACMIALGOTLYGVOKRITDVLGRPRVLPIDIRH"
868..2042
/gene="nifS"
868..2042
/gene="nifS"
/codon_start=1
/transl_table=11
/protein_id="CAA4848.1"
/db_xref="GI:297918"
/db_xref="SWISS-PROT:007177"
/translation="MTOPAIYLDNNATTVRPEVVAAMLPTFTEHFGNASSGHFGAO
AGLRKRAVAALIGASREBELFTSGTEANTATRSALAVODGERLITISAVEH
PAILLADLETEBEVRYHRTIPVDLGLDLDAATRAALSDRYAASTIMANNRDLPEV
POLAEBAHAGLHTDVAQAGLPLMSLADLEIDMSLSGKHGHPGVGALILRG
VAFRLKRGKQERGRAGTEVPAVLGMGLAEIVARMAADSRITATLEDWLESGI
IALGCCORIGDPAARLANTCTLAIFDSEAVLKLIAAGIAISGSCAGAPESP
VTRAAVPEFTAAHGVRFSLSDTTAELSHVLAILELPELAEHPVPTA"
2054..3202
/gene="nifV"
2054..3202
/gene="nifV"
/codon_start=1
/transl_table=11
/protein_id="CAA4848.1"
/db_xref="GI:297919"
/db_xref="SWISS-PROT:007179"
/translation="WTPDRTALCDTTLRDEQTAGVAFSLAERKATARALDRAGVAE
IEVGIAAGMAEVAETIRAVVAEIAHATFVWCRLRMDLDAOKTGKRVFAVPTST
AOLEGKLVDMDILRETAALVFCASDRGLGVGAEASRTDDPLRLAEVAAAG
AIRPFIADTIGLIDPLGAFRLVAELSAISLPIIMHAHNDGMAATNTIYAAHGAHP
LSVYVNGIGERAGNACCEVGALEAGSIDTGLDLCALPELSAVYAAASGRAPFAOKP
ITGDMFLAHESGHHVDAILKRDITIEDPCAPARFGEROIVIGKHSGLRAALL
EAGLPADDAVLDRLKPLRAHAHVRRVENGDLRRVVARLARPSPRT"
3232..3570
/gene="nifW"
3232..3570
/gene="nifW"
/codon_start=1
/transl_table=11
/protein_id="CAA4848.1"
/db_xref="GI:297920"
/db_xref="SWISS-PROT:007180"
/translation="WTPESPILALITLDSAEELFAFLGVPEIRVUNSSRLHIIMRF
GAVIREDMGTLEDIGFERARDAALRAADFAVSTPLKKEVERVETEAARRARV
GLEFLVKYKS"

BASE COUNT 570 a 1242 c 1191 g 567 t
ORIGIN
Query Match 4.4% Score 39.4; DB 1; Length 3570;
Best Local Similarity 50.8%; Pred. No. 78;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```


gene
/product="CONSERVED HYPOTHELICAL PROTEIN"
/protein_id="CAD15213.1"
/db_xref="GI:17428527"
/translation="MAQLTEAQHQLLLLDLDEQEARLHRLQTELESVPADAEVAEHR
YEEVDLADLEASERADAMRNHYRTELQAIIVARERLADGRGLCVDCGEAIPFLRLQ
APATACCACACORRERRWA"
complement(2254..2532)
/gene="RSC1512"
/note="RS03792"
complement(2234..2532)
/gene="RSC1512"
/function="miscellaneous; unknown"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHELICAL PROTEIN"
/protein_id="CAD15214.1"
/db_xref="GI:17428528"
/translation="WKSIEPPGRPOYSATNLTLTCVANDGSPACAVTALEADHFF
GAKSYLQEDLLALTEHRDAEDMARALFKLTDSHNIVLRSGHRRFGM"
complement(2866..3321)
/gene="rnhA"
/note="RSC1513: RS03791"
complement(2866..3321)
/gene="rnhA"
/EC_number="3.1.26.4"
/function="macromolecule metabolism; macromolecule
degradation; degradation of rna"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE RIBONUCLEASE HI PROTEIN"
/protein_id="CAD15215.1"
/db_xref="GI:17428529"
/translation="WQEVTVSDGACKGNPGLGNGTVLVSGHEKELFGEAVTTNN
RHELVIAEAPALRKPCKRVKYYTDSQYVQGISBWLAGMKARGMKTADKKPVKNDLL
WKTLDIELVTVHEVSHWVKGHAGHPGNERADALANKGEIARQAIOGA"
complement(3365..4138)
/gene="RSC1514"
/note="RS03790"
complement(3365..4138)
/gene="RSC1514"
/function="miscellaneous; hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHELICAL PROTEIN"
/protein_id="CAD15216.1"
/db_xref="GI:17428530"
/translation="MEGFHALOIGHPGFALRGALRENRPLIARAVYDIDDPAPIDTLPDA
DDPDGENDAEPPVQROATPRVYICRDIDELFPASOSIDLVALPHVLEFTDDPEHVEE
VARVLMPGRGLVITGFNPLSLMGHKGRRIGTSPFLPASQMTAFTRLADWLKLF
DITRERFCQCPNPTDKWLTQRTAFMEKAGDRMWIFGAVYMTQAVKRVRSVRLVSPA
WTKTRSPVLVAGCPVATPPSGSHTTKTPPVAAARDKSGA"
4227..5081
/gene="g10b"
/note="RSC1515: RS03789"
4227..5081
/gene="g10b"

Query Match	Similarity	4.4%	Score 39.4	DB 1	Length 202050
Best Local	Similarity	46.7%	Pred. No. 73		
Matches 161	Conservative	0	Mismatches 181	Indels 3	Gaps 1
QY 280	aaaccggaagacatgaaacaatgttcaaaagcgtacacgctgtggaacacgtgcaagcag	339			
Db 12802	AACCGGAACCGTCGTGACACGACATACGACACACTCCGACACCGCTTACTTGAACGGCGTG	12861			
QY 340	gctttggaacaaagcgcaagcgctgctgtlcatcacgcgcacacatcgcagctacgatttg	399			
Db 12862	ACGCTGGAAGACGTGCTCGAGATGCTGCACAAAGAAAAGCCGGTGGCGTATGTCGAG	12921			
QY 400	ggcggaagctatcatagccagcagtttcglttccaccctggaaccgcacatgtlaaagccgcg	459			
Db 12922	TACGGCGGCCAGACCCGCTGTGAAGTTCGCGCTCGACCTGTGAAGCCA---ACGGCGTGGCC	12978			
QY 460	aaatcaaaagcatagacaaaatcatctgcaagcggaagctgtgcgcgcgaagaagcaaac	519			
Db 12979	ATTCATCGGCACGACCGCGGACTATATTCACACGGCGCCGAAGACGCGAGCGCTTCCAGAG	13028			
QY 520	gcgcaccacgcgcatacaagggttcaaaacatlcataagccctgcgcgagcgcaagca	579			
Db 13039	CTGCTGCACGACCTGGCGCTCGGCCACGCCCAACCCACACCGCGCGCGGCGAAGACGAA	13098			
QY 580	accatcatcctgcgcgcagcaacgctccctctcgcgaaggaagcggc	624			
Db 13099	GCCCTCAAGCTTGCGCGACGAGATCGGCTTACCCGCTGGTGTCGCGC	13143			

SC7A1	32039 bp	DNA	linear	BCT 15-DEC-1998
LOCUS				
DEFINITION	Streptomyces coelicolor cosmid 7A1.			
ACCESSION	AI034447			
VERSION	AI034447.1			
KEYWORDS	GI:4007715			
	50S ribosomal protein L28; 50S ribosomal protein L32; acyl-acylphosphatase; acyltransferase; ammonium transporter; amt; ATP-dependent DNA helicase; chromosome associated protein; D-alanine-D-alanine ligase; dlla; formamidopyrimidine-DNA glycosylase; fpg; ftsy; glnB; glycerol-3-phosphate dehydrogenase; gpdA; KDO transferase; kdsB; methylase; nitrogen regulatory protein P-II; phosphomethylpyrimidine kinase; prokaryotic docking protein; regg; ribonuclease III; rnc; rnmB; rnmI; SRP54; sugar transporter; thiamine monophosphate kinase; thnD; thnL; transcriptional regulator.			
SOURCE	Streptomyces coelicolor A3(2).			
ORGANISM	Streptomyces coelicolor A3(2)			
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
REFERENCE	1 (bases 1 to 32039)			
AUTHORS	Murphy,L. and Harris,D.			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 32039)			
AUTHORS	Bentley,S.D., Parkhill,J., Barrell,B.G. and Randsdram,M.A.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (09-DEC-1998) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB0 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK			
REFERENCE	3 (bases 1 to 32039)			
AUTHORS	Reddenbach,M., Kleiser,H.M., Denapante,D., Eichner,A., Cullum,J., Kinasli,H. and Hopwood,D.A.			
JOURNAL	A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome			
MEDLINE	MOL. Microbiol. 21 (1), 77-96 (1996)			
COMMENT	97000351			
	Notes:			
	Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.			
	Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.			

FEATURES

SOURCE

misc_feature

gene

CDS

gene

CDS

(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg. SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 7A1 lies between 1C2 and 2E1 on the AseI-B genomic restriction fragment.

location/Qualifiers

1..32039

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid 7A1"

1..112

/note="Overlap with Streptomyces coelicolor cosmid 1C2 (EMBL:AL031124)."

/complement(97..735)

/gene="SC7A1.01c"

/complement(97..735)

/note="SC7A1.01c"

/note="SC7A1.01c, unknown, len: 212aa; similar to TR:P95112 (EMBL:283018) hypothetical protein from Mycobacterium tuberculosis (214 aa) fasta scores: opt: 420, z-score: 422.3, E(): 3.1e-16, (43.5% identity in 207 aa overlap) and TR:Q33128 (EMBL:299263) hypothetical protein from Mycobacterium leprae (216 aa) fasta scores: opt: 386, z-score: 414.0, E(): 9.1e-16, (40.6% identity in 207 aa overlap)."

/codon_start=1

/transl_table=1

/product="hypothetical protein SC7A1.01c"

/protein_id="CAA22400.1"

/db_xref="GI:4007716"

/translation="MONTLVYVYKALAAKASRLSDTADGIRPGIATAFADTYAALACRAVADVAVTDDARKGRELALAGAGVADPPGGGLAALHGAAYRAKRPSPVAALNDLPALRAELARVLAATQFPRAFLPDAAGIGTLLTVAPGDELAPFGADSRARHRAAGVELRDVADSVRDVDTGDLRSALALGVGPRTAVAARLIAGQ"

879..1673

/gene="SC7A1.02"

879..1673

/note="SC7A1.02"

/note="SC7A1.02, possible acyltransferase, len: 264aa; similar to several proposed acyltransferases eg. "TR:007809 (EMBL:297188) acyltransferase from Mycobacterium tuberculosis (251 aa) fasta scores: opt: 266, z-score: 384.5, E(): 4e-14, (29.7% identity in 229 aa overlap)."

/codon_start=1

/transl_table=1

/product="putative acyltransferase"

/protein_id="CAA22401.1"

/db_xref="GI:4007717"

/translation="MPRRRIQGFWRILAAVIGKPPVLVLIKRDGMENIPAEQGFITAVNHSVDPFAVIAHYQINTGVRPFLAKSGLFKKGFGAARNGTGQIPVYESTDAL

gene

CDS

misc_feature

misc_feature

gene

CDS

misc_feature

gene

CDS

AFRAIDAVERGECVAFYPEGTLTRDDPGMPEAKTGAARVALQKCPVIVAGMGNELIPPVAKKSPVLRKTHOYVAGPPVQLSPYRDEMTTEVLEKATEVIMAVTROLEIRCKAPETPPRRERIRERQRRRQAKSQVAPRRTRGPAEGST"

1670..2680

/gene="SC7A1.03"

1670..2680

/note="SC7A1.03"

/note="SC7A1.03, gpdA, glycerol-3-phosphate dehydrogenase, len: 366aa; similar to many eg. SW:GPD_BACSU glycerol-3-phosphate dehydrogenase from Bacillus subtilis (345 aa) fasta scores: opt: 874, z-score: 1221.4, E(): 0, (41.8% identity in 330 aa overlap). Contains a possible PS00017 ATP/GTP-binding site motif A (P-loop) and P1am match to entry PF01210 NAD_Gly3P_dh, NAD-dependent glycerol-3-phosphate dehydrogenase, score 295.10, E-value 8.8e-85."

/codon_start=1

/transl_table=1

/product="glycerol-3-phosphate dehydrogenase"

/protein_id="CAA22402.1"

/db_xref="GI:4007718"

/translation="MSKPKKAVFGTSGMGTAFGLVADACEVTLNGRAALDAVNSTRNPVLPGEVLEPENLRATTDAAEARQADPVLVPSQTLRAGLADPPLAPGTLVLSLMGVEIGSAMRSEVIGVAKGAREIAYVTGNLAREIAAMPAAVAVACDETVAQRQAQCHTYFRFYNTVDVAGCELGGAANVGLANGIDKGLGNAGSLITRGAEITRLGVALGADPLTFSGLAGLDVATCSPLSRNHTFTNKGKGLTEETNAVTKQTAEGVKSCSEVIDLARRHGVDMPITETVAIVHEKSPVAVAKELMSRAKPER"

1670..2647

/gene="SC7A1.03"

/note="P1am match to entry PF01210 NAD_Gly3P_dh, NAD-dependent glycerol-3-phosphate dehydrogenase, score 295.10, E-value 8.8e-85"

2600..2623

/gene="SC7A1.03"

/note="PS00017 ATP/GTP-binding site motif A (P-loop)." 2772..3941

/gene="SC7A1.04"

2772..3941

/note="SC7A1.04"

/note="SC7A1.04, ddla, D-alanine-D-alanine ligase, len: 389aa; similar to many eg. SW:DDLA_SALTY D-alanine-D-alanine ligase from Salmonella typhimurium (363 aa) fasta scores: opt: 742, z-score: 1018.5, E(): 0, (39.3% identity in 377 aa overlap). Contains PS00843 D-alanine-D-alanine ligase signature 1."

/codon_start=1

/transl_table=1

/product="D-alanine-D-alanine ligase"

/protein_id="CAA22403.1"

/db_xref="GI:4007719"

/translation="MSTENLPQNPQSPRRPRPRAVVGGRSSHGHSVYTAGAVLAAIDRTYDVLPIGIFTRDGMALTADEPMAALTERTPVEELAESEGVLLPVDPANREYVSEPGSPKALGVGAVFVPHGYGSDGTLOGLESGVYVAGVLAASVGOQREYKAVFTSYGLKGVGPAVIRPMDQSSGARKIYVDAGHGMPLEFKPARAGSSIGITKVDLGLDGLDAIEARHDKILVEALAKREIECVLEDEDPKRSVPAEIPPSSEHAYVDFEAKYIDSPGIVPAVLAEFAEVQRLAAVADAASCGLVRADFLTEDEFEVINEINTMGFTPISMYPMQMQASGVSYDELVDRLVQAALRPTGLR"

3162..3197

/gene="SC7A1.04"

/note="PS00843 D-alanine-D-alanine ligase signature 1."

complement(4119..4352)

/gene="SC7A1.05c"

complement(4119..4352)

/gene="SC7A1.05c"

/note="SC7A1.05c, conserved hypothetical protein, len: 77aa; similar to several putative/hypotheticals eg. TR:Q28868 (EMBL:AB01007) proposed transcriptional regulatory protein from Archaeoglobus fulgidus (77 aa) fasta scores: opt: 123, z-score: 217.3, E(): 8.3e-05, (31.9% identity in 69 aa overlap)." 3162..3197

/codon_start=1

/transl_table=1


```

repeat_region      /rpt_family="L1"
                    3015..3314
                    /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
                    3320..3519
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    3528..3591
                    /rpt_family="(TA)n"
repeat_region      /rpt_family="Alu"
                    3641..3716
                    /rpt_family="Alu"
repeat_region      /rpt_family="L1"
                    3719..4069
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    3925..3956
                    /rpt_family="AT-rich"
repeat_region      /rpt_family="AT-rich"
                    4070..4523
                    /rpt_family="ERV1"
misc_feature        /rpt_family="ERV1"
                    4307..4313
                    /note="match to EST AI185950 (NID:g3736588) qe29c07.s1"
repeat_region      4526..5529
                    /rpt_family="L1"
repeat_region      5530..5827
                    /rpt_family="L1"
                    /rpt_family="Alu"
                    5770..5783
                    /note="similar to Homo sapiens EST BG425880
                    (NID:g13332386)"
                    5828..5983
                    /rpt_family="L1"
repeat_region      5867..6004
                    /rpt_family="T-rich"
                    5984..6283
                    /rpt_family="Alu"
repeat_region      6284..6685
                    /rpt_family="L1"
                    6686..6988
                    /rpt_family="Alu"
repeat_region      6966..6988
                    /rpt_family="Alu"
                    6989..7344
                    /rpt_family="(A)n"
repeat_region      7344..7395
                    /rpt_family="L1"
                    7386..8002
                    /rpt_family="L1"
repeat_region      8003..8130
                    /rpt_family="L1"
                    8134..8624
                    /rpt_family="Alu"
repeat_region      8625..8660
                    /rpt_family="L1"
                    8661..8958
                    /rpt_family="Alu"
repeat_region      8959..9202
                    /rpt_family="L1"
                    9203..9495
                    /rpt_family="L1"
                    9486..10786
                    /rpt_family="Alu"
repeat_region      /rpt_family="L1"
                    10999..11033
                    /rpt_family="(TTG)n"
                    11078..11263
                    /rpt_family="L1"
repeat_region      11276..11303
                    /rpt_family="(TTTG)n"
                    11285..11600
                    /rpt_family="Alu"
repeat_region      11604..11813
                    /rpt_family="L1"
                    11814..12110
                    /rpt_family="Alu"
repeat_region      12111..12573
                    /rpt_family="L1"
                    12395..12933
                    /note="match to EST BG471643 (NID:g13403918)"
                    12480..12481
                    misc_feature

```

```

misc_feature        /note="match to EST B1752627 (NID:g15744205)"
                    12625..12829
                    /note="similar to Homo sapiens EST B115055
                    (NID:g14565956)"
misc_feature        /note="similar to Homo sapiens EST BG340632
                    12625..13125
                    (NID:g13147070)"
misc_feature        /note="similar to Homo sapiens EST BG421520
                    12625..12944
                    (NID:g13328026)"
                    12625..13039
                    /note="similar to Homo sapiens EST AW250630
                    (NID:g6593623)"
                    12625..12928
                    misc_feature

Query Match      4.4%; Score 39.2; DB 9; Length 93403;
Best Local Similarity 46.4%; Pred. No. 82;
Matches 128; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 358 gggctgtcttcatcaccgcgcacatcgccagctacgttggcgagcgtacatcagc 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55830 gggcgccctccggccctccggccctccggccctccggccctccggccctccggcc 55889
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 418 cagcagcttcgttcacacgtacacgtacacgtacacgtacacgtacacgtacacgtac 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55890 TCCTCGAAGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 55949
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 478 aaatcatcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55950 AGCTCTTGCACTCCCGCGAGCGGCTCTCGGGGTACAGCACCCTCGCAGGCC 56009
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 538 ggggtcaacaatcataaagccctgctgctgctgctgctgctgctgctgctgctgctg 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56010 AGCGCGAGCCCGCCACACAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 56069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 598 cagtccttcctccgaggaagcgagcgagcgagcgagcgagcgagcgagcgagcgag 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56070 ANGCTTCCCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCGCTCCCGCG 56105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 41
AC027550 133814 bp DNA linear HTG 09-JUN-2001
AC027550
LOCUS Homo sapiens chromosome 15 clone RP11-34M23 map 15, WORKING DRAFT
DEFINITION
ACCESSION AC027550.2 GI:7677918
VERSION AC027550.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 133814)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Homo sapiens chromosome 15, clone RP11-34M23.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 133814)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choquet,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,U., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,G., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheters,R.,
Meldrim,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,L., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

```

TITLE
JOURNAL
COMMENT

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 2, 2000 this sequence version replaced g1:7342295.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: 34_M23

Center clone name: L8928

----- Summary Statistics -----

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 121005 bases at least Q40

Consensus quality: 127719 bases at least Q30

Consensus quality: 129965 bases at least Q20

Insert size: 154000; agarose-fp

Insert size: 13114; sum-of-contigs

Quality coverage: 3.5 in Q20 bases; agarose-fp

Quality coverage: 4.1 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1037: contig of 1037 bp in length
* 1038 1137: gap of 100 bp
* 1138 2302: contig of 1165 bp in length
* 2303 2402: gap of 100 bp
* 2403 3543: contig of 1141 bp in length
* 3544 3643: gap of 100 bp
* 3644 5816: contig of 2173 bp in length
* 5817 5916: gap of 100 bp
* 5917 8729: contig of 2813 bp in length
* 8730 8829: gap of 100 bp
* 8830 10450: contig of 1621 bp in length
* 10451 10550: gap of 100 bp
* 10551 12885: contig of 2335 bp in length
* 12886 12985: gap of 100 bp
* 12986 14315: contig of 1330 bp in length
* 14316 14415: gap of 100 bp
* 14416 16318: contig of 1903 bp in length
* 16319 16418: gap of 100 bp
* 16419 19648: contig of 3230 bp in length
* 19649 19748: gap of 100 bp
* 19749 24321: contig of 4573 bp in length
* 24322 24421: gap of 100 bp
* 24422 28469: contig of 4048 bp in length
* 28470 28569: gap of 100 bp
* 28570 31830: contig of 3261 bp in length
* 31831 31930: gap of 100 bp
* 31931 33765: contig of 1835 bp in length
* 33766 33865: gap of 100 bp
* 33866 38584: contig of 4719 bp in length
* 38585 38684: gap of 100 bp
* 38685 42792: contig of 4100 bp in length
* 42793 42892: gap of 100 bp
* 42893 48877: contig of 5985 bp in length

FEATURES
source

* 48878 48977: gap of 100 bp
* 48978 53084: contig of 4107 bp in length
* 53085 53184: gap of 100 bp
* 53185 58248: contig of 5064 bp in length
* 58249 58348: gap of 100 bp
* 58349 65014: contig of 6666 bp in length
* 65015 65114: gap of 100 bp
* 65115 71933: contig of 6819 bp in length
* 71934 72033: gap of 100 bp
* 72034 78451: contig of 6418 bp in length
* 78452 78551: gap of 100 bp
* 78552 84834: contig of 6283 bp in length
* 84835 84934: gap of 100 bp
* 84935 91353: contig of 6419 bp in length
* 91354 91453: gap of 100 bp
* 91454 98364: contig of 6911 bp in length
* 98365 98464: gap of 100 bp
* 98465 108890: contig of 10426 bp in length
* 108891 108990: gap of 100 bp
* 108991 120523: contig of 11533 bp in length
* 120524 120623: gap of 100 bp
* 120624 133814: contig of 13391 bp in length.
Location/Qualifiers
1. 133814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-34M23"
/clone_lib="RPC1-11 Human Male BAC"
1. 1037
/note="assembly_fragment"
1138. 2302
/note="assembly_fragment"
2403. 3543
/note="assembly_fragment"
3644. 5816
/note="assembly_fragment"
5917. 8729
/note="assembly_fragment"
8830. 10450
/note="assembly_fragment"
10451. 10550
/note="assembly_fragment"
10551. 12885
/note="assembly_fragment"
12886. 14315
/note="assembly_fragment"
14316. 16318
/note="assembly_fragment"
16319. 16418
/note="assembly_fragment"
16419. 19648
/note="assembly_fragment"
19649. 24321
/note="assembly_fragment"
24322. 28469
/note="assembly_fragment"
28470. 31830
/note="assembly_fragment"
31831. 33765
/note="assembly_fragment"
33766. 33865
/note="assembly_fragment"
33866. 38584
/note="assembly_fragment"
38585. 42792
/note="assembly_fragment"
42793. 48877
/note="assembly_fragment"
48878. 53084
/note="assembly_fragment"
53085. 58248
/note="assembly_fragment"
58249. 58348
/note="assembly_fragment"

```

misc_feature /note="assembly_fragment"
55115..71933
misc_feature /note="assembly_fragment"
72034..78451
misc_feature /note="assembly_fragment"
78552..84834
misc_feature /note="assembly_fragment"
84935..91353
misc_feature /note="assembly_fragment"
91454..98364
misc_feature /note="assembly_fragment"
98465..108890
misc_feature /note="assembly_fragment"
108991..120523
misc_feature /note="assembly_fragment"
120624..133814
misc_feature /note="assembly_fragment"
31054 c 31211 g 34402 t 2704 others
BASE COUNT 34443 a
ORIGIN

```

```

Query Match 4.4%; Score 39.2; DB 2; Length 133814;
Best Local Similarity 46.4%; Pred. No. 81; Mismatches 148; Indels 0; Gaps 0;
Matches 128; Conservative 0;

```

```

358 gggcgtgtgtatcaacgcccacatcgagctacattggcgagcgtacatcaagc 417
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87971 GGGCGGCTCCCGCGCCGCCCGCCAGATCGACCTCCACAGCGCGCGTACCACTGC 88030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 cagcagcttcgttcacacgtacccatgtacaagccgcgaataacatgaagac 477
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88031 TCTCGAGAGCGCCGCGCTGCTGTGAGAGAAATGCGCGGAGAGCTCCAGAGCGCTCC 88090
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
478 aaatcatgacagggcgaggtgctgcgcaaaagcaaacgcccacgcgcgcatacaa 537
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88091 AGCTCTTGCACTCCCGCGACGGCTCTCTCTGGGGGTACAGCATCCCGCTGCAGCGCC 88150
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
538 ggggtcaacaacatcaagccctgcgagcgagcgaggaacacatcatctgcccagc 597
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88151 AGCCGCGAGCGCCACACAGCGCCGCGCGCCGCGCGCCGCGCGCGCGCGCC 88210
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 cagctcccttcctgcaggaagcgagcggtgtg 633
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88211 ATGCTTCGGGTCCCGCTCGCGCCGCTCGCGG 88246
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 42
AE004483/c 10202 bp DNA linear BCT 30-AUG-2000

LOCUS AE004483
DEFINITION genome.
ACCESSION AE004483 AE004091
VERSION AE004483.1 GI:9946313
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa.
Pseudomonas aeruginosa.
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.

REFERENCE
AUTHORS
1 (bases 1 to 10202)
Stover,C.K., Pham,X.-Q., Erwin,A.L., Mizoguchi,S.D., Warrenner,P.,
Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M.,
Gardner,R.L., Goltry,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y.,
Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R.,
Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen
Nature 406 (6799), 959-964 (2000)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
2 (bases 1 to 10202)
Stover,C.K., Pham,X.-Q., Erwin,A.L., Mizoguchi,S.D., Warrenner,P.,
Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J.,
Lagrou,M., Gardner,R.L., Goltry,L., Tolentino,E.,
Westbrock-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N.,

FEATURES source

Polger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H.,
Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Saeier,M.H.,
Hancock,R.E.M., Lory,S. and Olson,M.V.
Direct Submission
Submitted (16-MAY-2000) Department of Medicine and Genetics,
University of Washington Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
Location/Qualifiers
1..10202
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
complement(76..2277)
/gene="PA0454"
complement(76..2277)
/gene="PA0454"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC03843.1"
/db_xref="GI:9946314"
/translation="MEPSLTHTTRRLMALDKRAYSLSRYPRIALAGALLICWMDRIEL
LIPFLGILIASALTEDDNNWGLRLQALLVTLCPVAAAVELLFPYPIPLPSGMLSL
AFGMILIGALGERAAIASATLILSTYTMGVQROGEVADVHPELLIVAGAAVGL
LSVIMNLFNSNOPOOSLSRLPELGEYIKSLLEPLEKOLDVEGORLALANONKV
VVALNOKERTIINRLGHGRRPVRYSRIKIFYAODVHERASSSHPYNRLAEPPHS
DVAFRCORLUNOQCKACQALAPALRIROPDVPADRELAEDLOASLEHROOSNPMK
GLRSLCALANLTTLDRLKLAGSNPDALDEDSALIDEGORLALANONKV
TSLFRHGLRMALALITAGVYLWHTPTGYSWILLTVYCOPNYCATRLRLVORTIG
TVVGLVGMALIDLFPSPVLOSLEVAAGVAFPAFRSRYTLATAITLLVLCFMOQ
GDGYGLPLRFEDFLGSLIAGLAVELVDPMOGRRLNRLMANTLACNSYVLMQV
YANGKRDLDVETARRNAHNDALSTLSNMILLEPFRKDAIEIGREFLYSEHLLLS
YLSLGHNRSLPEPDAISGLIERAERLSLEFIASLSLSEKRPVAYISEEALAKE
LEQMPEDMDSDHRLVOTQLAHICQPLALPRTIAHLKRLAKAAGPLGQP"
complement(2358..3914)
/gene="db"
/note="PA0455"
complement(2538..3914)
/gene="db"
/codon_start=1
/transl_table=11
/product="RNA helicase Dbpa"
/protein_id="AAC03844.1"
/db_xref="GI:9946315"
/translation="MTSTAFSSLPISADLLANLDSLIGRENTPIQASLPILLIQGHDL
IAOKTSGSKTAARGLALISPLNRYEGCOALVYCPRRELADQVAKIRRLARADNI
KVLPLCGSVPRGYOASLEHGAHVVGPGRIORHLKRLTVLDGLNTLVLDADNML
DMGFYDSIAETIGQLPERROTLLSATYPPGCTEKLAKTEKMPDPROVESLHDSOLE
ORFEIDPKQRMADVRLLOHFRPQSCVAFRCQTRQCCQELADAKKORISALAHGL
EQRERDQVLTVEFANSCNVIVATVYAAGLIDALEVINVELADEPVIHRTGBCG
RAGEKGLSLVAPAEASRAQALIDLOKQPLAWLQULKASPEPLLPYSTLDIAGG
RKDKVRPGDLGALTGDAGLPGSAVGIKIFDFOAFVAIERGLAHQALKRLECKING
RSFKVRLI"
4262..4471
/gene="PA0456"
4262..4471
/gene="PA0456"
/codon_start=1
/transl_table=11
/product="probable cold-shock protein"
/protein_id="AAC03845.1"
/db_xref="GI:9946316"
/translation="MSRQNGYVKNWNETKGYGFTTPESGPVYVHFRATIEGKTLA
EGOKVSFEVVGOKMQAERVQVIN"
4622..5146
/gene="PA0457"
4622..5146
/gene="PA0457"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC03846.1"
/db_xref="GI:9946317"

gene
CDS
/translation="MKPVQSRRLAFALLSLSSSSASCHATEHMGIMQSWLMPERM
QAKAGLDTGALTSLDARINIRYKDKGERVARDVLPETADSKAPVSVFERVRLI
KVGAGGSDSRPVYAMDICGAKTLREQFSLRDRGMNVPVLLGRTLEHLGAVDVR
TTRKPTCSALAO"
complement(5516..6949)
/gene="PA0458"
complement(5516..6949)
/gene="PA0458"
/codon_start=1
/transl_table=11
/product="probable MFS transporter"
/protein_id="AAG03867.1"
/db_xref="GI:9946318"

gene
CDS
/translation="MSEPFQLEPRRIARQLEPMVAFAEMQALDGLTINTALPSMASL
NENPLRMQAVIAYLLTVALIPASGMIDRESTRFVGLAVLFSGLSLACLSL
ELVGARIVHVGEGALMMPYRVLIVYRQDVLVRFVITLSPVITGLLPLAGTIGG
LVEASVHWIIFELNIPGLIGLVAKMKIDPLSPSPRSDSGILLREGSWSLISA
LEGLGELHSLHREVVLLIGVLITAYIRAIRIKPPEPSLFARFTAVGILGNL
FARLGSALDFPLPILQVGLIPSTAGMTMPLALFAVAAVPAKPLIDFEGYRKL
LVGNLILGLIAGFSLVQDTPYVLLHLSLIGAVNSLOFTAMNTLLIDLODSNA
SSGNSLMSVYVQLSISLVACAALLGEGFTDLDGTQDVLGAFHATYVSVGLMSLA
AAIFQIGSEDDGRTSPRRVDPKDPDL"
7570..10122
/gene="PA0459"
7570..10122
/gene="PA0459"
/codon_start=1
/transl_table=11
/product="probable ClpA/B protease ATP binding subunit"
/protein_id="AAG03848.1"
/db_xref="GI:9946319"

BASE COUNT 1814 a 3301 c 3579 g 1508 t
ORIGIN
Query Match 4.4%; Score 39; DB 1; Length 10202;
Best Local Similarity 48.8%; Pred. No. 95;
Matches 105; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 381 catgagcagctagcatttggcgagcagctacacgacgacgttcggttcaacctgac 440
DB 5906 CTTGGCGCTGCTGACACGAGACACCCCTGCTGCTGCTGCTGACCTGAGCCTGCT 5847
QY 441 cgcacatcaagcgcgcgaataacgacatgacgaataatcatgagcgagcgaggt 500
DB 5846 CGGCGCGGTGAACCTCCCTGCAATTCACGGCGGATGACACCTTGCATCTGCA 5787
QY 501 ggcgcgcgaacgaacacgcgcgcacgcgcgcacgcgcgcacgcgcgcacgcgcgc 560
DB 5786 GGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 5727
QY 561 cctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 595
DB 5726 CCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5692

RESULT 43
AE001885/c

LOCUS AE001885 14790 bp DNA linear BCT 22-NOV-1999
DEFINITION Deinococcus radiodurans R1 section 22 of 229 of the complete
chromosome 1
ACCESSION AE001885 AE000513
VERSION AE001885.1 GI:6457902
KEYWORDS
SOURCE Deinococcus radiodurans
ORGANISM Deinococcus radiodurans
Bacteria: Thermus/Deinococcus group; Deinococcales; Deinococcaceae;
Deinococcus.
1 (bases 1 to 14790)
White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D.,
Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L.,
Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
Vamathevan, J. J., Lam, P., McDonald, L., Uterback, T., Zaleski, C.,
Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M., et al.
Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1
Science 286 (5444), 1571-1577 (1999)
2 (bases 1 to 14790)
White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D.,
Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L.,
Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
Vamathevan, J. J., Lam, P., McDonald, L., Uterback, T., Zaleski, C.,
Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M., et al.
Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1
Science 286 (5444), 1571-1577 (1999)

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
1 (bases 1 to 14790)
White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D.,
Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L.,
Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
Vamathevan, J. J., Lam, P., McDonald, L., Uterback, T., Zaleski, C.,
Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M., et al.
Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1
Science 286 (5444), 1571-1577 (1999)
2 (bases 1 to 14790)
White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D.,
Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L.,
Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
Vamathevan, J. J., Lam, P., McDonald, L., Uterback, T., Zaleski, C.,
Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M., et al.
Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1
Science 286 (5444), 1571-1577 (1999)

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
1. 14790
/organism="Deinococcus radiodurans"
/strain="R1"
/db_xref="taxon:1299"
/chromosome="1"
78..1205
/gene="DR0229"
78..1205
/note="similar to GB:AE000782 percent identity: 63.16;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="endoglucanase, putative"
/protein_id="AAF09817.1"
/db_xref="GI:6457903"
/translation="MERVAVLGRHPRNSRNSRLHNSDPPSPINRDLFELLSQAP
SGAERRADVMKKEATFARVSEHYSVYAELEPENGPAIAGMDELGLYSHD
DKGHSVYGVGMDPOVYVGOIRILAGGDLGVGKKAHWAEEERKASLEND
WIDGSLDADEVKAKVPGCTGVTGCGPMVNNKVSRLDRVAFYVLRRLKRY
ELKRIYAVGTISOEIGVFGAOTSIRKIDPVAGIADVTHTETDOPGVSEKRYGVAPG
SGANLSVGPMTNPTLRQLTEAAGRENTPYLSASGRYGTDAATLVIRAGVPTGV
SIPNRYMSPSEBMDERDVAKCIDILAMIRGLEETLEFR"
1234..1719
/gene="DR0230"
1234..1719
/note="similar to PID:1652545 percent identity: 61.43;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF09818.1"
/db_xref="GI:6457904"

gene
CDS
1712..2188
/gene="DR0231"
1712..2188

```

/gene="DR0231"
/note="similar to SP:P33899 PID:466744 GB:U00096
PID:1790034 percent identity: 65.28; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="RNA methyltransferase, putative"
/protein_id="AA09819.1"
/db_xref="GI:6457905"
/translation="MTDAKSGPLLVLPPEKAGNNGVATCSVLGAELHTRP
GFHDHREERAVMDYIOGVTLHEIYKWSDFQADPAKVRFSSTHAEVHTRAEK
RGDYLLFEPESRILPWLRDALPKLPQGRGRSLNSVAAGAAFEAGQIERW"
2281..2508
/gene="DR0232"
/note="DR0232"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA09830.1"
/db_xref="GI:6457916"
/translation="MSENRQKANKYQRIISGLLNAERDVRLARAGTDEARAKANV
RLETLRALEIYASRKLAYGERPPEERT"
2505..2702
/gene="DR0233"
/note="DR0233"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, ArsR family"
/protein_id="AA09820.1"
/db_xref="GI:6457906"
/translation="MSADLARAALFRALSHPARLLRLATWALESGEHLANLVOLA
PAIVSHLAQLEVGMLTVQ"
2746..3075
/gene="DR0234"
/note="DR0234"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA09831.1"
/db_xref="GI:6457917"
/translation="MSLTSIVQGSADVPFVEDYEQRVLRAPLRGKLIQIPQRKR
DVVLRFLAGLFEPRRYPEAEVNLIGEYHDFEFTLRRLVAGLDERDGETWREP
TEPSSL"
3081..4766
/gene="DR0235"
/note="DR0235"
/codon_start=1
/transl_table=11
/product="6-aminohexanoate-cyclic-dimer hydrolase"
/protein_id="AA09821.1"
/db_xref="GI:6457907"
/translation="MTRKMSFLASGASTHEONGEMSGRCCSDARVIRRTALSCIM
LTFAYDRDLALDLQFLRGLSAEDMCTAAIHRQVNVVAVNAVYPLAYDGLAQ
RATDARARQATGAPVAGVFLYKDGSLAGVPHGTGRAYRDQIPENDDELVRM
QAALPLIKGTNPPEFALMGVTEPELHGPRRNWDIGRTGSGSGSASVAVGIYPL
AGADGSGSTRIPASCCGLRGRKRVCCGVCYCPMGCAVENVHITFSVDSAL
LDLEQGPDAAGALPLSPERPYSEVEKEPGRURIGSTAHPLGRSVHPCVAVGA
ARLEISGHEVEEVALPMDGPALAQAFIMLVFGETGSLAALDITLGRPARADVAV
TWLLGOLGRYSADPAAARSNVHARAGRPHONDILLTFVLAATPPLQIELOPR
GVQALRAAQQMDVSGILRRSGVDLADLDIEKMPYQTLANTLQOPMSVLIHMTA
DGLPVGQVFAPIARDEVILRIAGOLEQARPWARRRTVGRRAVGEATYSGSA"

```

```

4763..5998
/gene="DR0236"
/note="DR0236"
/codon_start=1
/transl_table=11
/product="aminopeptidase"
/protein_id="AA09822.1"
/db_xref="GI:6437908"
/translation="MTPSSFDERLARRAELLVKVNLIPAGKILLVAPLAPVRL
VARAAYRAGALDVREYEDHEIALALYEDGSDAMGFLPWATEQVOHMDLDGVARLS
IRGDDPALKNVNERIARRSKLMAQASREVSAVSGFAVWVAAMSTPMAQAVP
ELPEAEVAGCLMDIDFEKTRADLDPDGAMNDHYAALARRSAHLNDKQYAAVHLEL
GTDLVGLAEERHIOGGAEEAKNGICVGPVNPTEDEFTMPHRDRVNGVAVASRPLSR
GOVEGIVRERENGKAYVSAKEGEDTLKSLIARDDECAHILGELIAYVAAPAVQCT
LEFNTLFDEPNASHIALGRCTPTNVQGADQQAAGAGNDSLIHVDMIGTPEASVD
GITADGQREPLRMGEWMI"
6169..6750
/gene="DR0237"
/note="DR0237"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl cis-trans isomerase,
cyclophilin-type"
/protein_id="AA09823.1"
/db_xref="GI:6437909"
/translation="MRLMSDYNAEGFTQTPYLSTERTKFTQAPELGDAIEPKQYR
AVLETSKRIVELYPDEAPMTVNSFALLHNNHYDGIKFRIVIDGMAQGDPTG
MGPGYKFEDEAGNHRSGKGVLSMANAPGTNGSQFFITFATPDLNHRHVFQK
VVEGIDVLDITRIQIPQMGSTPDIETAVLEK"
6888..8146
/gene="DR0238"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
GB:AL009126 percent identity: 52.56; identified by
sequence similarity; putative; conserved hypothetical
protein; authentic frameshift"
8146..8550
/gene="DR0239"
/note="DR0239"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA09832.1"
/db_xref="GI:6457918"
/translation="MRWVPVLLPLLLSLACODREARQNAELARRVALESEVSLK

```

Query Match 4.4%; Score 39; DB 1; Length 14790;
Best Local Similarity 48.8%; Pred. No. 94; Mismatches 110; Indels 0; Gaps 0;
Matches 105; Conservative 0;

QY	293	tcgaacaatgttcaagcggtacacgctggaacagctgacgacgcttggaacaag	352
DB	11805	TGAGAGACGGTGGCCGCCGCCGCTACGACGACGCGGCTGGAAGGCGCTCAAG	11746
QY	353	gcgaagggctgtgttcacacgcgcacatcggaagctagattggcggaagctaca	412
DB	11745	CCGAGCTCGCGCTGCTCCTCGCGCAAGATCACCCGCGCTGCTGCTCAAG	11686
QY	413	tcggcagcgagcttcggttcacacctgacgcgcatggaacagcgcaaatcaagga	472
DB	11685	GCTTCAAGGCGAGCGGTTCAATCAAGCTCGCCTGTGGGCGGAGCGTGGCAAGACG	11626
QY	473	tagacaaatcatgcagcgagcgaggttcgagcg	507

Db 11625 TGGTCATCACCATCGAGAAGTTCACCGACTCGAC 11591

```

/OD_Xref="GI:14329741"
/translaton="MAKRLVFGIIVIALVLTAAEGEASRLQCERELQESSLEACR
QVVDQDLACRLPRSTGLQMRCCQLRDVYSAKRHVAVSQVARYEQGTAVVPPKGSITP

```

